GRAPHICAL KREIN SIGNATURE THEORY AND EVANS-KREIN FUNCTIONS

RICHARD KOLLÁR* AND PETER D. MILLER†

Abstract. Two concepts, evidently very different in nature, have proved to be useful in analytical and numerical studies of spectral stability in nonlinear wave theory: (i) the Krein signature of an eigenvalue, a quantity usually defined in terms of the relative orientation of certain subspaces that is capable of detecting the structural instability of imaginary eigenvalues and hence their potential for moving into the right half-plane leading to dynamical instability under perturbation of the system, and (ii) the Evans function, an analytic function detecting the location of eigenvalues. One might expect these two concepts to be related, but unfortunately examples demonstrate that there is no way in general to deduce the Krein signature of an eigenvalue from the Evans function, for example by studying derivatives of the latter.

The purpose of this paper is to recall and popularize a simple graphical interpretation of the Krein signature well-known in the spectral theory of polynomial operator pencils. Once established, this interpretation avoids altogether the need to view the Krein signature in terms of root subspaces and their relation to indefinite quadratic forms. To demonstrate the utility of this graphical interpretation of the Krein signature, we use it to define a simple generalization of the Evans function — the Evans-Krein function — that allows the calculation of Krein signatures in a way that is easy to incorporate into existing Evans function evaluation codes at virtually no additional computational cost. The graphical interpretation of the Krein signature also enables us to give elegant proofs of index theorems for linearized Hamiltonians in the finite dimensional setting: a general result implying as a corollary the generalized Vakhitov-Kolokolov criterion (or Grillakis-Shatah-Strauss criterion) and a count of real eigenvalues for linearized Hamiltonian systems in canonical form. These applications demonstrate how the simple graphical nature of the Krein signature may be easily exploited.

Key words. Krein signature, Evans function, operator pencils, stability, Hamiltonian systems

AMS subject classifications. 37K45, 47A75, 15A18, 15A22, 47A10, 47J10, 35B35, 65L07

1. Introduction.

This paper concerns relations among several concepts that are commonly regarded as useful in the analysis of (generalized) eigenvalue problems such as those that occur in the stability theory of nonlinear waves: *Krein signatures* of eigenvalues, *Evans functions*, and *index theorems* (also known as *inertia laws* or *eigenvalue counts*) governing the spectra of pairs of related operators. The most elementary setting in which many of these notions appear is the stability analysis of equilibria for finite-dimensional Hamiltonian systems, and we take this opportunity right at the beginning of the paper to introduce the key ideas in this simple setting, including a beautiful and simple graphical method of analysis.

1.1. A graphical method for linearized Hamiltonians.

The most important features of the linearized dynamics of a finite dimensional Hamiltonian system close to a critical point are determined by the values of ν in the *spectrum* $\sigma(JL)$ of the (nonselfadjoint) problem

$$JLu = \nu u\,, \qquad\qquad L^* := \overline{L^{\mathsf{T}}} = L,\ J^* = -J. \tag{1.1}$$

*Department of Applied Mathematics and Statistics, Faculty of Mathematics, Physics, and Informatics, Mlynská dolina, 842 48 Bratislava, Slovakia (kollar@fmph.uniba.sk), supported by National Science Foundation under grant DMS-0705563, and by the European Commission Marie Curie International Reintegration Grant 239429.

†Department of Mathematics, University of Michigan, 530 Church St, Ann Arbor, MI 48109-1043 (millerpd@umich.edu), supported by National Science Foundation under grants DMS-0807653 and DMS-1206131.

Here J is an invertible skew-Hermitian matrix and L a Hermitian matrix of the same dimension (both over the complex numbers $\mathbb{C}$).[1] The conditions on J require that the dimension of the space be even and hence L and J have dimension $2n \times 2n$. Indeed, a Hamiltonian system linearized about an equilibrium takes the form

$$\frac{dy}{dt} = JLy, \tag{1.2}$$

where y denotes the displacement from equilibrium. The spectral problem (1.1) arises by looking for solutions growing as $e^{\nu t}$ by making the substitution $y(t) = e^{\nu t}u$, and we say that (1.2) is *spectrally stable* if $\sigma(JL)$ consists of only purely imaginary numbers. The points $\nu \in \sigma(JL)$ for which $\mathrm{Re}\{\nu\} \neq 0$ are called the *unstable spectrum* [2] of (1.2). The *linearized energy* associated with the equation (1.2) is simply the quadratic form $E[u] := (Lu, u)$, and the fundamental conservation law corresponding to (1.2) is that $dE[y]/dt = 0$ on all solutions $y = y(t)$.

With the use of information on the spectrum $\sigma(L) \subset \mathbb{R}$ of L our goal is to characterize the spectrum $\sigma(JL) \subset \mathbb{C}$ of JL and in particular to determine (i) the part of $\sigma(JL)$ in the open right half of the complex plane (i.e., the unstable spectrum) and (ii) the potential for purely imaginary points in $\sigma(JL)$ to collide on the imaginary axis under suitable perturbations of the matrices resulting in bifurcations of Hamiltonian-Hopf type [55, 58] in which points of $\sigma(JL)$ leave the imaginary axis and further destabilize the system.

The key idea is to relate the purely imaginary spectrum of (1.1) to invertibility of the selfadjoint (Hermitian) linear pencil $\mathcal{L}(\lambda)$ (see Section 2 for a proper definition of operator pencils) defined by

$$\mathcal{L} = \mathcal{L}(\lambda) := L - \lambda K, \qquad K := (iJ)^{-1} = K^*, \quad \lambda := i\nu \in \mathbb{R}. \tag{1.3}$$

Indeed, upon multiplying (1.1) by $-iJ^{-1}$ one sees that (1.1) is equivalent to the generalized spectral problem

$$Lu = \lambda K u\,. \tag{1.4}$$

The graphical method [23, 24, 49, 57] is based on the observation that the imaginary spectrum of (1.1), or equivalently the real spectrum λ of (1.4), can be found by studying the spectrum of the eigenvalue pencil $\mathcal{L}(\lambda)$ and the way it depends on $\lambda \in \mathbb{R}$, i.e., by solving the selfadjoint eigenvalue problem

$$\mathcal{L}(\lambda)u(\lambda) = (L - \lambda K)u(\lambda) = \mu(\lambda)u(\lambda) \tag{1.5}$$

parametrized by $\lambda \in \mathbb{R}$ (see Fig. 1.1). Clearly, $\nu = -i\lambda$ is a purely imaginary number in $\sigma(JL)$ if and only if $\lambda \in \mathbb{R}$ is a value for which $\mathcal{L}(\lambda)$ has a nontrivial kernel, that is, $0 \in \sigma(\mathcal{L}(\lambda))$. It is also easy to see from the right-hand plot in Fig. 1.1 that the particular values λ for which $0 \in \sigma(\mathcal{L}(\lambda))$ correspond in a one-to-one fashion to

[1] On notation: we use $\overline{\nu}$ to denote the complex conjugate of a complex number ν, while A^* denotes the conjugate transpose of A, or more generally when an inner product is understood, the adjoint operator of A. We use (u, v) to denote an inner product on vectors u and v, linear in u and conjugate-linear in v.

[2] Although the points ν with $\mathrm{Re}\{\nu\} < 0$ correspond to decaying solutions of (1.2), due to the basic Hamiltonian symmetry of (1.1) they always go hand-in-hand with points $-\overline{\nu}$ in the right half-plane that are associated with exponentially growing solutions of (1.2), explaining why they are also included in the unstable spectrum.

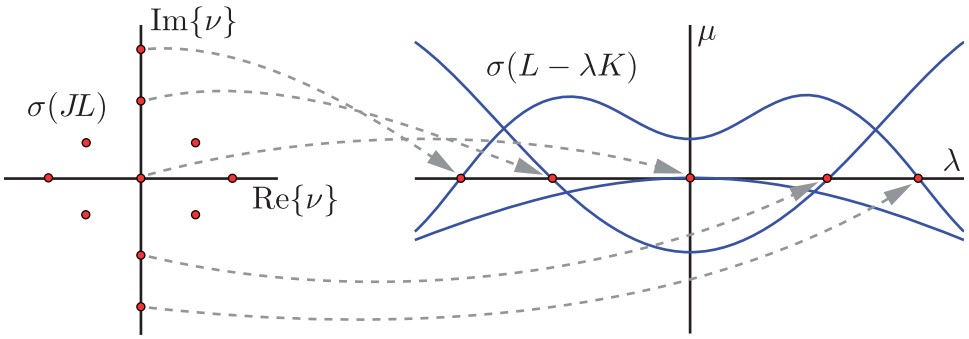

FIG. 1.1. *The key to the graphical method lies in the correspondence of purely imaginary points ν in the spectrum of JL with zeros of spectrum of the pencil $\mathcal{L}(\lambda) = L - \lambda K$, where $K := (iJ)^{-1}$.*

intercepts of curves $\mu = \mu(\lambda) \in \sigma(\mathcal{L}(\lambda))$ with the axis $\mu = 0$. This identification also holds true for spectrum of higher multiplicity (for details see Theorem 3.1).

To make a clear distinction between the two types of spectrum that are related by the above simple argument we now introduce the following terminology. The points λ in $i\sigma(JL)$ will be called *characteristic values* of $\mathcal{L}$ with corresponding invariant *root spaces* spanned by *generalized characteristic vectors* or *root vectors*. Therefore a purely imaginary point in $\sigma(JL)$ corresponds to a real characteristic value of $\mathcal{L}$, but the characteristic values of $\mathcal{L}$ are generally complex (see Definition 2.1). On the other hand, given an arbitrary real number λ, the points μ in $\sigma(\mathcal{L}(\lambda))$ will be called *eigenvalues* of $\mathcal{L}(\lambda)$ with corresponding *eigenspaces* spanned by (genuine, due to selfadjointness as $\lambda \in \mathbb{R}$) *eigenvectors*. When we consider how $\mu = \mu(\lambda)$ depends on $\lambda \in \mathbb{R}$ we will frequently call $\mu(\lambda)$ an *eigenvalue branch*.

1.2. Use of the graphical method to find unstable spectrum. Purely imaginary points in $\sigma(JL)$ correspond in a one-to-one fashion to real characteristic values of $\mathcal{L}$ simply by rotation of the complex plane by 90°. The most obvious advantage of the graphical method is that the presence and location of real characteristic values of $\mathcal{L}$ can be easily read off from plots of the eigenvalue branches $\mu(\lambda)$ of $\mathcal{L}(\lambda)$ using the following elementary observations. Selfadjointness of L and skewadjointness of J imply that $\sigma(JL)$ consists of pairs $(\nu, -\overline{\nu})$, and hence the characteristic values of $\mathcal{L}$ come in complex-conjugate pairs. Indeed, if $u \in \mathbb{C}^{2n}$ is a (right) root vector of JL corresponding to $\nu \in \sigma(JL)$, then $(J^{-1}u)^*$ is a left root vector of JL corresponding to $-\overline{\nu} \in \sigma(JL)$. Therefore the JL-invariant subspace spanned by root spaces corresponding to non-real characteristic values of $\mathcal{L}$ is even-dimensional. Since the base space $X = \mathbb{C}^{2n}$ is also even-dimensional, the number of real characteristic values of $\mathcal{L}$ (counting multiplicities) is even, and consequently the total number of intercepts of eigenvalue branches $\mu(\lambda)$ with $\mu = 0$ (again, counting multiplicities) must be even.

If furthermore J and L are real matrices (see Definition 5.4 for a more general definition of reality in less obvious contexts) then $\sigma(JL)$ also consists of pairs $(\nu, \overline{\nu})$ and we say that (1.1) has *full Hamiltonian symmetry*. In such a case, $\sigma(\mathcal{L}(\lambda))$ is necessarily symmetric with respect to the vertical ($\lambda = 0$) axis, so the plots of eigenvalue branches μ as functions of λ are left-right symmetric. This is the case illustrated in Fig. 1.1.

Finally, the dimension of the problem limits the maximal number of intercepts of all branches $\mu(\lambda)$ with the axis $\mu = 0$ to $2n$. This is also true for intercepts with every

horizontal axis $\mu = \mu_0$, as one can consider the problem $J(L-\mu_0 I)u = \nu u$. With this basic approach in hand, we now turn to several elementary examples.

$\lhd$ *Example 1. Definite Hamiltonians.* It is well-known that if L is a definite (positive or negative) matrix, then $\sigma(JL)$ is purely imaginary and nonzero. Indeed, if $u \in \mathbb{C}^{2n}$ is a root vector of JL corresponding to $\nu \in \sigma(JL)$, then $Lu = \nu J^{-1}u$, so taking the Euclidean inner product with u gives

$$0 \neq (Lu, u) = \nu(J^{-1}u, u),$$

and hence neither ν nor $(J^{-1}u, u)$ can be zero. Moreover, $(Lu, u) \in \mathbb{R}$ and $(J^{-1}u, u) \in i\mathbb{R}$, and hence $\nu = (Lu, u)/(J^{-1}u, u)$ is a purely imaginary nonzero number.

This simple fact can also be deduced from a plot of the eigenvalue branches $\mu(\lambda)$ of the selfadjoint pencil $\mathcal{L}(\lambda)$. Let us assume without loss of generality that L is positive definite. We need just three facts:

- The $2n$ eigenvalue branches $\mu(\lambda)$ may be taken to be (we say this only because ambiguity can arise in defining the branches if $\mathcal{L}(\lambda)$ has a nonsimple, but necessarily semisimple, eigenvalue for some $\lambda \in \mathbb{R}$) continuous functions of λ. In fact, they may be chosen to be *analytic* functions of λ, although we do not require smoothness of any sort here.
- The $2n$ eigenvalue branches $\mu(\lambda)$ are all positive at $\lambda = 0$ since $\mathcal{L}(0) = L$.
- By simple perturbation theory, $\sigma(\mathcal{L}(\lambda)) = -\lambda\sigma(K) + O(1)$ as $|\lambda| \to \infty$. Since K is Hermitian and invertible, it has $m \leq 2n$ strictly positive eigenvalues and $2n - m$ strictly negative eigenvalues. Hence there are exactly m eigenvalue branches $\mu(\lambda)$ tending to $-\infty$ as $\lambda \to +\infty$, and exactly $2n - m$ eigenvalue branches $\mu(\lambda)$ tending to $-\infty$ as $\lambda \to -\infty$.

That all $2n$ characteristic values of $\mathcal{L}$ are nonzero real numbers, and hence $\sigma(JL) \subset i\mathbb{R}$ making the system (1.2) spectrally stable, therefore appears as a simple consequence of the intermediate value theorem; m branches $\mu(\lambda)$ necessarily cross $\mu = 0$ for $\lambda > 0$ and $2n-m$ branches cross for $\lambda < 0$. Since $m+(2n-m) = 2n$ exhausts the dimension of $X = \mathbb{C}^{2n}$, all characteristic values have been obtained in this way. This approach provides the additional information that exactly m of the points in $\sigma(JL)$ are negative imaginary numbers. Note that in the case of full Hamiltonian symmetry, $m = n$.

To illustrate this phenomenon, the branches $\mu(\lambda)$ corresponding to the following specific choices of L positive definite and JL having full Hamiltonian symmetry:

$$L = \frac{1}{2}\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & 4 \end{pmatrix}, \quad J = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ -1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{pmatrix},$$

are plotted in the left-hand panel of Fig. 1.2. $\rhd$

$\lhd$ *Example 2. Indefinite Hamiltonians and instability.* If L is indefinite, then $\sigma(JL)$ is not automatically confined to the imaginary axis. As an illustration of this phenomenon, consider the matrices

$$L = \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -2 \end{pmatrix}, \quad J = \begin{pmatrix} 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 1 \\ -2 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{pmatrix},$$

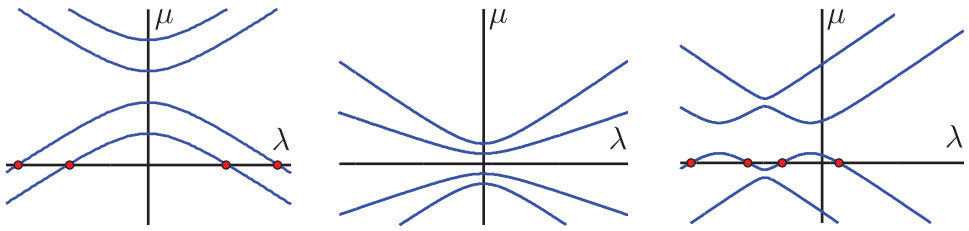

FIG. 1.2. *The four eigenvalue branches $\mu(\lambda)$ of $\mathcal{L}(\lambda)$ in Examples 1–3. The real (characteristic) values λ for which eigenvalues $\mu(\lambda)$ of $\mathcal{L}(\lambda)$ intersect the horizontal line $\mu = 0$ and that correspond to purely imaginary points $\nu = -i\lambda \in \sigma(JL)$ are indicated with red dots.*

which again provide full Hamiltonian symmetry, but now L has two positive and two negative eigenvalues. The corresponding eigenvalue branches $\mu(\lambda)$ are plotted against $\lambda \in \mathbb{R}$ in the middle panel of Fig. 1.2. Obviously $\sigma(\mathcal{L}(\lambda))$ is bounded away from zero as $\lambda \in \mathbb{R}$ varies, and hence there are no real characteristic values of $\mathcal{L}$, implying that $\sigma(JL) \cap i\mathbb{R} = \emptyset$. The full Hamiltonian symmetry in this example implies that $\sigma(JL)$ is organized in quadruplets. Thus either $\sigma(JL)$ consists of two positive and two negative real points, or $\sigma(JL)$ consists of a single quadruplet of non-real and non-imaginary points. Hence all of the spectrum is unstable and the graphical method easily predicts spectral instability in this case. ▷

◁ *Example 3. Stability with indefinite Hamiltonians.* If J and L are not both real the full Hamiltonian symmetry and hence the left-right symmetry of the union of eigenvalue branches $\mu(\lambda)$ is broken. In the right-hand panel of Fig. 1.2 one such example is shown corresponding to

$$L = \frac{1}{2}\begin{pmatrix} 2 & -8 & 5 & 0 \\ -8 & -3 & 0 & 0 \\ 5 & 0 & 9 & -8 \\ 0 & 0 & -8 & 13 \end{pmatrix}, \quad J = -i\begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}. \tag{1.6}$$

Here we see one eigenvalue branch $\mu(\lambda)$ of $\mathcal{L}(\lambda)$ intersecting $\mu = 0$ transversely at four nonzero locations. Since $\dim(X) = 4$, this implies $\sigma(JL) \subset i\mathbb{R} \setminus \{0\}$, and the graphical method predicts spectral stability. Note that L is indefinite in this case showing that indefiniteness of L does not imply the existence of unstable spectrum. That is, while definiteness of L implies stability, the converse is false. ▷

1.3. Use of the graphical method to obtain Krein signatures. These three examples have illustrated the use of the graphical method to count purely imaginary points in $\sigma(JL)$, which are encoded as zero intercepts of the eigenvalue curves $\mu = \mu(\lambda)$. We now wish to dispel any impression that the location of the intercepts might be the only information contained in plots like those in the right-hand panel of Fig. 1.1 and in Fig. 1.2. This requires that we introduce briefly the notion of the Krein signature of a purely imaginary point $\nu \in \sigma(JL) \cap i\mathbb{R}$.

Krein signature theory [2, 3, 45, 46] allows one to understand aspects of the dynamics of Hamiltonian flows near equilibria (see [58] for a recent review). Equilibria that are local extrema of the associated linearized energy are energetically stable, and this is the situation described in Example 1 above. However, even in cases when an

equilibrium is not a local extremum, the energetic argument can still predict stability if the linearized energy is definite on an invariant subspace; the fact is that many Hamiltonian systems of physical interest have constants of motion independent of the energy (e.g. momentum) and this means that effectively the linearized dynamics should be considered on a proper invariant subspace of the linearized phase space corresponding to holding the constants of motion fixed. Therefore, in such a situation only certain subspaces are relevant for the linearized energy quadratic form, and definiteness of the form can be recovered under an appropriate symplectic reduction. Of course, for (1.2), the invariant subspaces are simply the root spaces of JL, and for the particular case of genuine characteristic vectors u corresponding to $\nu \in \sigma(JL)$, it is easy to see how the definiteness of the linearized energy relates to stability. Indeed, if $JLu = \nu u$, then for invertible J we have $Lu = \nu J^{-1}u = i\nu Ku$ and by taking the inner product with u and using selfadjointness of L and K one obtains two identities:

$$\operatorname{Re}\{\nu\}(Ku, u) = 0 \quad \text{and} \quad (Lu, u) = -\operatorname{Im}\{\nu\}(Ku, u). \tag{1.7}$$

If ν lies in the unstable spectrum, then $\operatorname{Re}\{\nu\} \neq 0$ and the first equation requires that $(Ku, u) = 0$ which from the second equation implies $E[u] = (Lu, u) = 0$, so the linearized energy is indefinite on the subspace. On the other hand, if ν is purely imaginary, then the first equation is trivially satisfied but the second gives no information about (Lu, u). This calculation suggests that (Lu, u), or equivalently (Ku, u), carries nontrivial information when $\nu \in \sigma(JL) \cap i\mathbb{R}$. We will define[3] the *Krein signature* of a simple purely imaginary point $\nu \in \sigma(JL)$ (or equivalently of the purely real characteristic value $\lambda = i\nu$ of $\mathcal{L}$) as follows:

$$\kappa(\lambda) := -\operatorname{sign}(Ku, u), \quad \text{where } \lambda \in \mathbb{R},\ JLu = -i\lambda u,\ K = (iJ)^{-1}, \text{ and } u \neq 0. \tag{1.8}$$

For simple $\nu \in i\mathbb{R}$, (Ku, u) is necessarily nonzero[4] by the invertibility of J.

Consider a simple point $\nu = -i\lambda$ in $\sigma(JL) \cap i\mathbb{R}$. If the matrices J and L are subjected to sufficiently small admissible perturbations, then (i) ν remains purely imaginary and simple and (ii) $\kappa(\lambda)$ remains constant; as an integer-valued continuous function (of J and L), the only way a Krein signature $\kappa(\lambda)$ can change under perturbation is if $\nu = -i\lambda$ collides with another nonzero purely imaginary eigenvalue. From this point of view, one of the most important properties of the Krein signature is that it captures the susceptibility of a point $\nu \in \sigma(JL) \cap i\mathbb{R}$ to *Hamiltonian-Hopf bifurcation* [3, 71, 74] (see also MacKay [55] for a geometric interpretation of the Krein signature within this context) in which two simple purely imaginary points of $\sigma(JL)$ collide under and leave the imaginary axis. Indeed, for bifurcation to occur, it is necessary that the colliding points have opposite Krein signatures. In fact, this condition

[3]Unfortunately, it is equally common in the field for the Krein signature to be defined as $\operatorname{sign}(Lu, u)$, and while the latter is more obviously connected to the linearized energy, our definition is essentially equivalent according to (1.7) and provides a more direct generalization beyond the context of linear pencils considered here.

[4]The argument is as follows: since $\nu = -i\lambda$ is a simple point of $\sigma(JL)$, the space X can be decomposed as $X = \operatorname{span}\{u\} \oplus X_c$ where X_c is the complementary subspace invariant under JL, and $JL + i\lambda$ is invertible on X_c. Therefore, $v \in X_c$ can be represented in the form $v = (JL + i\lambda)w$ for some $w \in X_c$, and it follows that for all $v \in X_c$ we have

$$(J^{-1}u, v) = (J^{-1}u, (JL + i\lambda)w) = -((LJ + i\overline{\lambda})J^{-1}u, w) = -(J^{-1}(JL + i\overline{\lambda})u, w).$$

Since $\lambda \in \mathbb{R}$ and $JLu = -i\lambda u$ we see that $(J^{-1}u, \cdot)$ vanishes on X_c. But since J is invertible this form cannot vanish on all of X and hence it must be definite on $X \ominus X_c$ implying that $(J^{-1}u, u) \neq 0$. See [44] for more details.

is also sufficient, in the sense that if satisfied there exists an admissible deformation of JL that causes the Hamiltonian-Hopf bifurcation to occur. On the other hand, imaginary points of $\sigma(JL)$ with the same Krein signatures cannot leave the imaginary axis even if they collide under perturbation.

The definition (1.8) makes the Krein signature $\kappa(\lambda)$ appear as either an algebraic quantity (computed via inner products), or possibly a geometric quantity (measuring relative orientation of root spaces with respect to the positive and negative definite subspaces of the linearized energy form). We would now like to emphasize a third interpretation of the formula, related to the graphical method for linearized Hamiltonians introduced above. Indeed, from the point of view of Krein signature theory, the main advantage of the reformulation of (1.1) as (1.5) is that if ν_0 is a purely imaginary simple point in $\sigma(JL)$ corresponding to the intersection of an eigenvalue branch $\mu(\lambda)$ with $\mu = 0$ at the real characteristic value $\lambda = \lambda_0 = i\nu_0$ of $\mathcal{L}$, then the Krein signature $\kappa(\lambda_0)$ turns out to have a simple interpretation as the slope of the branch at the intersection point:

$$\kappa(\lambda_0) = \operatorname{sign}\left[\frac{d\mu}{d\lambda}(\lambda_0)\right]. \tag{1.9}$$

Therefore, without computing any inner products at all, we can read off the Krein signatures of the imaginary spectrum of JL in Examples 1 and 3 above simply from looking at the diagrams in the corresponding panels of Fig. 1.2. For Example 1, the signatures of the negative characteristic values are $\kappa = 1$ while those of the positive characteristic values are $\kappa = -1$, and the only possibility for the system to become structurally unstable to Hamiltonian-Hopf bifurcation would be for a pair of characteristic values to collide at $\lambda = 0$ (forced by full Hamiltonian symmetry), and this clearly requires L to become indefinite under perturbation. On the other hand, Example 3 represents a somewhat more structurally unstable case; in order of increasing λ the signatures are $\kappa = 1, -1, 1, -1$. Therefore all of the pairs of adjacent real characteristic values are susceptible to Hamiltonian-Hopf bifurcation.

The formula (1.9) yields perhaps the easiest proof that simple real characteristic values of the same Krein signature cannot undergo Hamiltonian-Hopf bifurcation even should they collide; locally one has two branches with the same direction of slope, and the persistent reality of the roots as the branches evolve can be seen as a consequence of the Intermediate Value Theorem. To prove (1.9), one differentiates (1.5) with respect to λ at $\lambda = \lambda_0$ and $\mu = 0$, obtaining the equation $(L - \lambda_0 K)u'(\lambda_0) = (\mu'(\lambda_0) + K)u$. Taking the inner product with the characteristic vector $u = u(\lambda_0)$ satisfying $(L - \lambda_0 K)u = 0$ and using selfadjointness of $L - \lambda_0 K$ for $\lambda_0 \in \mathbb{R}$ gives

$$\mu'(\lambda_0)(u, u) = -(Ku, u). \tag{1.10}$$

Since $(u, u) > 0$, the expression (1.9) immediately follows from the definition (1.8). This shows that not only are the locations of the intercepts of the curves $\mu = \mu(\lambda)$ important, but it is also useful to observe the way the curves cross the $\mu = 0$ axis.

1.4. Generalizations. Organization of the paper. The notion of Krein signature of real characteristic values has become increasingly important in the recent literature in the subjects of nonlinear waves, oscillation theory, and integrable systems [5, 7, 8, 10, 29, 36, 41, 43, 44, 72]. In many of these applications, the situation is more general than the one we have considered so far. One direction in which the theory can be usefully generalized is to replace the matrix $\mathcal{L} = L - \lambda K$ with a more general

Hermitian matrix function of a real variable λ resulting in a *matrix pencil* that is generally nonlinear in λ. Another desirable generalization is to be able to work in infinite-dimensional spaces where the operators involved are, say, differential operators as might occur in wave dynamics problems. The basic graphical method described above can also be applied in these more general settings.

The main applications we have in mind are for linearized Hamiltonian systems of the sort that arise in spectral stability analysis of nonlinear waves, i.e., linear eigenvalue problems. However many of the ideas used here can be traced back to a nearly disjoint but well-developed body of literature concerning nonlinear eigenvalue pencils and matrix polynomials [21, 23, 57]. The thought to connect aspects of this theory to problems in stability of nonlinear waves appears to have come up quite recently, although similar ideas were already used in [49]. In §2 we review the basic theory of matrix and operator pencils, which lays the groundwork for both the generalization to nonlinear dependence on the characteristic value and the generalization to infinite-dimensional spaces. Then, in §3 we give a precise definition of Krein signature along the lines of (1.8) and show how also for operator pencils there is a way to deduce the Krein signature from the way that an eigenvalue curve passes through $\mu = 0$, a procedure that is a direct generalization of the alternate formula (1.9).

In §4 we consider the problem of relating Krein signatures to a common tool used to detect eigenvalues, the so-called *Evans function* [1, 18, 64]. Unfortunately, attempts to deduce the Krein signature of eigenvalues from properties of the Evans function itself are easily seen to be inconclusive at best. However, the graphical (or perhaps topological) interpretation of the Krein signature as in (1.9) suggests a simple way to modify the traditional definition of the Evans function in such a way that the Krein signatures are all captured. This modification is even more striking when one realizes that the Evans function itself is based on a (different) topological concept [1], a Mel'nikov-like transversal intersection between stable and unstable manifolds. We call this modification of the Evans function the *Evans-Krein function*, and we describe it also in §4. The main idea is that, while in the linearization of Hamiltonian systems the Evans function restricted to the imaginary ν-axis characterizes the *product* of individual algebraic root factors, the Evans-Krein function is able to separate these factors with the help of the additional parameter μ. The use of the Evans-Krein function therefore allows these different root factor branches to be traced numerically, without significant changes to existing Evans function evaluating codes.

In §5 we extend the kind of simple arguments used to determine spectral stability in Examples 1–3 above to give short and essentially topological proofs of some of the well-known index theorems for nonselfadjoint spectral problems that were originally proven by very different, algebraic methods [24, 34, 36, 37, 65]. To keep the exposition as simple as possible, we present our new proofs in the finite-dimensional setting. The added value of the graphical approach is that it makes the new proofs easy to visualize, and hence to remember and generalize. Finally, we conclude in §6 with a brief discussion of related open problems.

Our paper features many illustrative examples. Readers trying to understand the subject for the first time may find it useful to pay special attention to these.

2. Matrix and Operator Pencils.

2.1. Basic terminology and examples. In the literature the terms *operator pencil* or *operator family* frequently refer to the same type of object: a collection of linear operators depending on a complex parameter lying in an open set $S \subset \mathbb{C}$, that is, a map $\mathcal{L} = \mathcal{L}(\lambda)$ from $\lambda \in S$ into a suitable class of linear operators from one

Banach space, X, into another, Y. Perhaps the most common type of pencil is a so-called *polynomial pencil* for which $\mathcal{L}(\lambda)$ is simply a polynomial in λ with coefficients that are fixed linear operators. (The sub-case of a linear pencil has already been introduced in Section 1.1.) If X and Y are finite-dimensional Banach spaces, we have the special case of a *matrix pencil*. The coefficients of a polynomial matrix pencil $\mathcal{L}(\lambda)$ can (by choice of bases of X and Y) be represented as constant matrices of the same dimensions. We will only consider the case in which X and Y have the same dimension, in which case the coefficient matrices of a polynomial matrix pencil are all square. Specializing in a different direction, if $X = Y$ is a (self-dual) Hilbert space, a pencil (operator or matrix) is said to be *selfadjoint* if $S = \overline{S}$ and $\mathcal{L}(\overline{\lambda}) = \mathcal{L}(\lambda)^*$. With the choice of an appropriate orthonormal basis, the coefficients of a selfadjoint polynomial matrix pencil all become Hermitian matrices.

An operator pencil consisting of *bounded* linear operators $\mathcal{L}(\lambda)$ on a fixed (possibly infinite-dimensional) Banach space X is said to be *holomorphic* at $\lambda_0 \in S$ if there is a neighborhood D of λ_0 in which $\mathcal{L}(\lambda)$ can be expressed as a convergent (in operator norm) power series in λ (see also [57, pp. 55–56]). If D contains a real open interval $I \subset \mathbb{R} \cap S$, we say that $\mathcal{L}$ is *continuously differentiable* on I if the restriction of $\mathcal{L}$ to I is continuously norm-differentiable. For holomorphic (continuously differentiable) matrix pencils, the individual matrix entries of the matrix $\mathcal{L}(\lambda)$ are all holomorphic functions of λ near λ_0 (continuously differentiable functions on I).

The theory of operator pencils, and matrix pencils in particular, is well-developed in the literature. The review articles by Tisseur and Meerbergen [69] and Mehrmann and Voss [60] are general references that give a numerous applications of operator pencils and survey suitable numerical methods for their study. Polynomial pencils are particularly well-understood and have an extensive spectral theory [21, 23, 57]. While most of the theory is concerned with finite dimensional matrix pencils (called λ-*matrices* or *gyroscopic systems*) some results have been obtained for general operator pencils (see [47, 48, 57] and references therein). The spectral theory of λ-matrices under various conditions was developed in detail by Lancaster *et al.* [30, 50, 52, 53]; see also [51] for the related subject of perturbation theory of analytic matrix functions.

In order to better motivate the theory of linear and nonlinear operator pencils, we first give some concrete examples of how they arise naturally in several applications.

$\lhd$ *Example 4. Spectral problems in inverse-scattering theory for integrable partial differential equations.* As is well known, some of the most interesting and important nonlinear partial differential equations of mathematical physics including the *modified focusing nonlinear Schrödinger equation*

$$i\phi_t + \frac{1}{2}\phi_{xx} + |\phi|^2\phi + i\alpha(|\phi|^2\phi)_x = 0, \quad \alpha \in \mathbb{R} \tag{2.1}$$

governing the envelope $\phi(x,t)$ of (ultrashort, for $\alpha \neq 0$) pulses propagating in weakly nonlinear and strongly anomalously dispersive optical fibers, and the *sine-Gordon equation*

$$u_{tt} - u_{xx} + \sin(u) = 0 \tag{2.2}$$

arising in the analysis of denaturation of DNA molecules and the modeling of superconducting Josephson junctions, are *completely integrable systems*. One of the key implications of complete integrability is the existence of an *inverse-scattering transform* for solving the Cauchy initial-value problem for $x \in \mathbb{R}$ in which (2.1) is given the complex-valued Schwartz-class initial condition $\phi(x,0) = \phi_0(x)$ and (2.2) is given

the real-valued Schwartz-class[5] initial conditions $u(x,0) = f(x)$ and $u_t(x,0) = g(x)$, and in each case the solution is desired for $t > 0$. The inverse-scattering transform explicitly associates the initial data for each of these equations to a certain auxiliary linear equation involving a spectral parameter, the spectrum of which essentially encodes all of the key properties of the solution for $t > 0$. For the modified focusing nonlinear Schrödinger equation (2.1) with $\alpha \neq 0$ the auxiliary linear equation for the inverse-scattering transform is the so-called WKI spectral problem due to Wadati *et al.* [73] with spectral parameter ξ and vector-valued unknown $v = v(x)$:

$$\mathcal{L}_{\rm WKI}(\xi)v = 0, \quad \text{where} \quad \mathcal{L}_{\rm WKI}(\xi) := \xi^2 W_2 + \xi W_1 + W_0, \tag{2.3}$$

involving a quadratic pencil $\mathcal{L}_{\rm WKI}(\xi)$, with coefficients being the linear operators

$$W_2 := 2i\sigma_3, \quad W_1 := -2i\alpha \begin{pmatrix} 0 & \phi_0 \\ \overline{\phi}_0 & 0 \end{pmatrix}, \quad W_0 := \alpha \frac{d}{dx} - \frac{1}{2} i\sigma_3. \tag{2.4}$$

Here σ_3 is a Pauli spin matrix[6]. In the special case $\alpha = 0$, one has to use a different auxiliary linear equation known as the Zakharov-Shabat problem [76] which is commonly written in the form

$$\frac{dv}{dx} = -i\xi\sigma_3 v + \begin{pmatrix} 0 & \phi_0 \\ -\overline{\phi}_0 & 0 \end{pmatrix} v \tag{2.5}$$

which by multiplication on the left by $i\sigma_3$ is obviously equivalent to a usual eigenvalue problem for an operator that is non-selfadjoint with respect to the $L^2(\mathbb{R})$ inner product (augmented in the obvious Euclidean way for vector functions v). However, note that if one sets $\lambda = i\xi$, $v(x) = e^{i\pi\sigma_3/4} w(x)$, and multiplies through on the left by $i\sigma_1 e^{-i\pi\sigma_3/4}$, the Zakharov-Shabat problem is also equivalent to the equation

$$\mathcal{L}_{\rm ZS}(\lambda)w = 0, \quad \text{where} \quad \mathcal{L}_{\rm ZS}(\lambda) := \lambda Z_1 + Z_0 \tag{2.6}$$

involving a linear pencil $\mathcal{L}_{\rm ZS}(\lambda)$ with coefficients

$$Z_1 := \sigma_2, \quad Z_0 := i\sigma_1 \frac{d}{dx} - \begin{pmatrix} \overline{\phi}_0 & 0 \\ 0 & \phi_0 \end{pmatrix}. \tag{2.7}$$

On the other hand, the initial-value problem for the sine-Gordon equation (2.2) is solved by means of the Faddeev-Takhtajan problem:

$$\mathcal{L}_{\rm FT}(\xi)v = 0, \quad \mathcal{L}_{\rm FT} := \xi F_1 + F_0 + \xi^{-1} F_{-1}, \tag{2.8}$$

a problem for a rational operator pencil with coefficients

$$F_0 := 4i\frac{d}{dx} - g\sigma_2 \quad F_{\pm 1} := \sin(\tfrac{1}{2}f)\sigma_1 \mp \cos(\tfrac{1}{2}f)\sigma_3. \tag{2.9}$$

[5]More generally, one only assumes that $\sin(f(x))$ is Schwartz class to admit the physically interesting possibility of nonzero *topological charge* in the initial data in which the angle f increases by a nonzero integer multiple of 2π as x ranges over $\mathbb{R}$.

[6]The Pauli spin matrices are:

$$\sigma_1 := \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_2 := \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma_3 := \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

In all of these cases, the values of ξ for which there exists a nontrivial $L^2(\mathbb{R})$ solution v parametrize the amplitudes and velocities of the *soliton* components of the solution of the corresponding (nonlinear) initial-value problem. The solitons are localized coherent structures that appear in the long-time limit, and as such it is important to have accurate methods to determine the location of any such discrete spectrum. Bounds for the discrete spectrum of the WKI spectral problem can be found in [12], and very sharp results that under some natural qualitative conditions on the initial data confine the discrete spectrum to the imaginary axis for the Zakharov-Shabat problem and the unit circle for the Faddeev-Takhtajan problem were found by Klaus and Shaw [42] and by Bronski and Johnson [7] respectively. In particular, the techniques used in [7, 42] can be interpreted in terms of Krein signatures. Indeed, one of the key conditions required in [42] is that the initial condition $\phi_0(x)$ is a real-valued function, which makes $\mathcal{L}_{\mathrm{ZS}}(\lambda)$ a *selfadjoint* linear pencil. The significance of selfadjointness will become clear later. $\triangleright$

$\triangleleft$ *Example 5. Hydrodynamic stability.* Consider a steady plane-parallel shear flow of a fluid in a two-dimensional channel with horizontal coordinate $x \in \mathbb{R}$ and vertical coordinate $z \in (z_1, z_2)$, and let $U = U(z)$ be the horizontal velocity profile. If the stream function of such a flow is perturbed by a Fourier mode of the form $\Phi(z)e^{ik(x-ct)}$ with horizontal wavenumber $k \in \mathbb{R}$, then in the case of an inviscid fluid the dynamical stability of such a flow is governed by the *Rayleigh equation* [13]:

$$\mathcal{L}_{\mathrm{R}}(c)\Phi = 0, \quad \mathcal{L}_{\mathrm{R}}(c) := cL_1 + L_0, \tag{2.10}$$

where Φ is subjected to the boundary conditions $\Phi(z_1) = \Phi(z_2) = 0$, and where $\mathcal{L}_{\mathrm{R}}(c)$ is a linear operator pencil with coefficients

$$L_1 := \frac{d^2}{dz^2} - k^2, \quad L_0 := U'' - UL_1. \tag{2.11}$$

For a viscous fluid, the unperturbed flow is characterized by the Reynolds number R, and its stability is determined from the *Orr-Sommerfeld equation* [13]:

$$\mathcal{L}_{\mathrm{OS}}(c)\Phi = 0, \quad \mathcal{L}_{\mathrm{OS}}(c) := cM_1 + M_0, \tag{2.12}$$

where Φ is subjected to the "no-slip" boundary conditions $\Phi(z_1) = \Phi(z_2) = 0$ and $\Phi'(z_1) = \Phi'(z_2) = 0$, and where $\mathcal{L}_{\mathrm{OS}}(c)$ is a linear operator pencil with coefficients

$$M_1 := ikRL_1, \quad M_0 := ikRL_0 + L_1^2. \tag{2.13}$$

In both cases, the values of $c \in \mathbb{C}$ for which there exists a nontrivial solution Φ are associated with exponential growth rates of $-ikc$, and hence the flow is stable to perturbations of wavenumber k if the corresponding values of c are all real. $\triangleright$

$\triangleleft$ *Example 6. Traveling wave stability in Klein-Gordon equations.* Let $V : \mathbb{R} \to \mathbb{R}$ be a potential function. The Klein-Gordon equation (of which the sine-Gordon equation (2.2) is a special case for $V(u) := 1 - \cos(u)$) is

$$u_{tt} - u_{xx} + V'(u) = 0. \tag{2.14}$$

Travelling wave solutions $u(x,t) = U(z)$, $z = x - ct$, of phase speed c satisfy the Newton-type equation $(c^2 - 1)U''(z) + V'(U(z)) = 0$. The wave is stationary in the co-moving frame in which (2.14) is rewritten for $u = u(z,t)$ in the form

$$u_{tt} - 2cu_{zt} + (c^2 - 1)u_{zz} + V'(u) = 0. \tag{2.15}$$

Writing $u = U(z) + v(z,t)$ and linearizing for v small one obtains

$$v_{tt} - 2cv_{zt} + (c^2 - 1)v_{zz} + V''(U(z))v = 0, \tag{2.16}$$

and seeking solutions of the form $v(z,t) = \phi(z)e^{i\lambda t}$ for $\lambda \in \mathbb{C}$ one arrives at a spectral problem involving a quadratic pencil:

$$\mathcal{L}_{\rm KG}(\lambda)\phi = 0, \quad \mathcal{L}_{\rm KG} := \lambda^2 L_2 + \lambda L_1 + L_0, \tag{2.17}$$

where the coefficients are

$$L_2 := 1, \quad L_1 := 2ic\frac{d}{dz}, \quad L_0 := (1-c^2)\frac{d^2}{dz^2} - V''(U(z)). \tag{2.18}$$

Note that $\mathcal{L}_{\rm KG}(\lambda)$ is an example of a (formally) selfadjoint quadratic operator pencil. Spectral stability is deduced [35] if all values of λ for which there exists a nontrivial solution ϕ are purely real. $\triangleright$

$\triangleleft$ *Example 7. The Rayleigh-Taylor problem.* Here we give an example of an operator pencil involving partial differential operators and nonlocality (through a divergence-free constraint). The Rayleigh-Taylor problem concerns the stability of a stationary vertically stratified incompressible viscous fluid of equilibrium density profile $\rho_e(z)$. Making a low Mach number approximation, assuming a small perturbation of the zero velocity field, and linearizing the Navier-Stokes equations, one is led to consider normal mode perturbations of the form $e^{\lambda t}\mathbf{u}(x,y,z)$ where the linearized velocity field satisfies appropriate no-slip boundary conditions and

$$-\lambda^2\rho_e\mathbf{u} + \nabla\cdot(\rho_e\mathbf{u})\mathbf{g} + \lambda(\nabla p - \eta\triangle\mathbf{u}) = 0\,, \qquad \nabla\cdot\mathbf{u} = 0, \tag{2.19}$$

where $\mathbf{g}$ is the gravitational acceleration field, $\eta > 0$ is the viscosity, and p is the pressure term needed to satisfy the incompressibility constraint [28]. A weak reformulation on an appropriate divergence-free space allows (2.19) to be cast into the form of an equivalent quadratic pencil with parameter λ on a Hilbert space. For non-Newtonian Maxwell linear viscoelastic fluids, the pencil that arises in the Rayleigh-Taylor problem is a cubic polynomial [43]. $\triangleright$

The preceding examples all involve polynomial operator pencils (or, like the rational pencil appearing in the Faddeev-Takhtajan spectral problem in Example 4, that can easily be converted into such). However, it is important to observe that spectral problems for non-polynomial operator pencils also occur very frequently in applications, especially those involving a mix of discrete symmetries and continuous symmetries for which the dispersion relation for linear waves is transcendental. A fundamental example is the following.

$\triangleleft$ *Example 8. Delay differential equations* Non-polynomial operator pencils appear naturally in systems of differential equations with delays [33]. Consider the system

$$\dot{x}(t) = Ax(t) + Bx(t-\tau), \tag{2.20}$$

for $x \in \mathbb{C}^n$, where $\tau > 0$ is a fixed delay and A, B are complex $n \times n$ matrices. To study the stability of solutions to (2.20) with exponential time-dependence of the form $x(t) = e^{\lambda t}x_0$ one needs to solve the spectral problem

$$\mathcal{L}_{\rm DDE}(\lambda)x_0 = 0, \tag{2.21}$$

where $\mathcal{L}_{\mathrm{DDE}}$ is the essentially transcendental matrix pencil

$$\mathcal{L}_{\mathrm{DDE}}(\lambda) := \lambda\mathbb{I} - A - e^{-\tau\lambda}B. \tag{2.22}$$

The existence of values of $\lambda \in \mathbb{C}$ with $\mathrm{Re}\{\lambda\} > 0$ for which there exists a nonzero solution x_0 of (2.21) indicates instability of the system (2.20). $\triangleright$

Other applications of operator pencils include the analysis of electric power systems [59], least-squares optimization, and vibration analysis [69]. Various examples of non-polynomial operator pencils are described in the documentation to the MATLAB Toolbox NLEVP [4], including applications to the design of optical fibers and radio-frequency gun cavity analysis. Non-polynomial operator pencils also arise in band structure calculations for photonic crystals [14], another example of a system in which discrete symmetry (entering through the lattice structure of the crystal) interacts with continuous symmetry (time translation). An application of quadratic operator pencils to second-order in time Hamiltonian equations can be found in [9], in which a particular case of the generalized "good" Boussinesq equation is studied. Finally, note that in an alternative to the approach to linearized Hamiltonians described in §1.1, the so-called *Krein matrix* method of Kapitula [36] relates a general linearized Hamiltonian spectral problem to a non-polynomial operator pencil.

2.2. Spectral Theory of Operator Pencils. In this subsection we summarize the theoretical background needed for our analysis. In addition to a proper definition of the spectrum (characteristic values) of an operator pencil $\mathcal{L}$ and its relation to the spectra $\sigma(\mathcal{L}(\lambda))$ of the individual operators $\mathcal{L}(\lambda)$ making up the pencil $\mathcal{L}$ as λ varies, our method relies on a kind of continuity of $\sigma(\mathcal{L}(\lambda))$ with respect to λ, which can be obtained with appropriate general assumptions. Note that even in the general case of operator pencils on infinite dimensional spaces, finite systems of eigenvalues have many properties similar to those of eigenvalues of matrices [39]. This allows us to study a wide class of problems, although infinite systems of eigenvalues may exhibit various types of singular behavior.

2.2.1. Matrix pencils. Matrix pencils and their perturbation theory are studied in [21, 22, 23, 39, 57] with a particular emphasis on polynomial matrix pencils. The finite dimensional setting allows a simple formulation of the spectral problem.

Definition 2.1 (Characteristic values of matrix pencils). *Let $\mathcal{L} = \mathcal{L}(\lambda)$ be a matrix pencil on $X = \mathbb{C}^N$ defined for $\lambda \in S \subset \mathbb{C}$. The characteristic values of $\mathcal{L}$ are the complex numbers $\lambda \in S$ that satisfy the characteristic equation* $\det(\mathcal{L}(\lambda)) = 0$. *The set of all characteristic values is called the spectrum of $\mathcal{L}$ and is denoted $\sigma(\mathcal{L})$.*

While the matrices involved all have size $N \times N$ for each $\lambda \in S$, the characteristic equation need not be a polynomial in λ and matrix pencils in general can have an infinite number of characteristic values. However, characteristic equations of polynomial matrix pencils of degree p in λ are polynomial of degree at most pN in λ, with equality if and only if the leading coefficient is an invertible matrix.

On the other hand, the eigenvalues μ of the related eigenvalue problem $\mathcal{L}(\lambda)u = \mu u$ satisfy the modified characteristic equation $\det(\mathcal{L}(\lambda) - \mu\mathbb{I}) = 0$ [39, II.2.1, p. 63] and belong to the λ-dependent spectrum $\sigma(\mathcal{L}(\lambda))$ (understood in the usual sense) of the $\mathcal{L}(\lambda)$. Of course, as μ solves a polynomial equation of degree N in μ, the total algebraic multiplicity of the eigenvalues μ is equal to $\dim(X) = N$, independent of $\lambda \in S$. It is well-known that if the matrix pencil $\mathcal{L}$ is *holomorphic* at $\lambda_0 \in S$ and if all eigenvalues μ of $\mathcal{L}(\lambda_0)$ are simple, then the N eigenvalue branches $\mu = \mu(\lambda)$ are locally holomorphic functions of λ near λ_0. In fact this is an elementary consequence of the Implicit

Function Theorem. The possible singularities at exceptional points λ (corresponding to non-simple eigenvalues μ) are well understood. For linear pencils of the form $\mathcal{L}(\lambda) = L_0 + \lambda L_1$ it is worth emphasizing that the individual eigenvalue functions $\mu = \mu(\lambda)$ are generally *not* linear functions of λ; moreover, the corresponding eigenprojections $P_{\mu(\lambda)}$ can have poles as functions of λ. See Motzkin and Taussky [61, 62] for a study of special conditions under which the eigenvalues of linear pencils are indeed linear in λ and the corresponding eigenprojections are entire functions. Furthermore, mere continuity of $\mathcal{L}(\lambda)$ does not imply continuity of eigenvectors (see [39, II.1.5, p.110] for specific examples). The situation simplifies for *selfadjoint* holomorphic matrix pencils as the following theorem indicates.

THEOREM 2.2 ([39, II.6.1, p. 120, Theorem 6.1]). *Let $\mathcal{L}(\lambda)$ be a selfadjoint holomorphic matrix pencil defined for $\lambda \in S = \overline{S}$ with $\mathbb{R} \subset S$. Then, for $\lambda \in \mathbb{R}$, the eigenvalue functions $\mu = \mu(\lambda)$ can be chosen to be holomorphic. Moreover, for $\lambda \in \mathbb{R}$, the (orthonormal) eigenvectors can be chosen as holomorphic functions of λ.*

Thus even if a k-fold eigenvalue μ occurs for some $\lambda_0 \in \mathbb{R}$, locally this can be viewed as the intersection of k real holomorphic branches $\mu = \mu(\lambda)$ (and similarly for the corresponding eigenvectors). Note that analyticity is not a necessary condition for mere differentiability of eigenvalues μ and eigenvectors for $\lambda \in \mathbb{R}$. Indeed, according to Rellich [68], selfadjoint continuously differentiable matrix pencils have continuously differentiable eigenvalue and eigenfunction branches [39, II.6.3, p. 122, Theorem 6.8].

2.2.2. Operator pencils. In passing to infinite-dimensional spaces, we want to restrict our attention to holomorphic pencils, and to handle unbounded (e.g. differential) operators, we need to generalize the definition given earlier. A convenient generalization is the following.

DEFINITION 2.3 (Holomorphic families of type (A) [39, VII.2.1, p. 375]). *An operator pencil $\mathcal{L}$, $\mathcal{L}(\lambda) : D(\lambda) \subset X \to X$, defined for $\lambda \in S$ is called a holomorphic family of type (A) if the domain $D = D(\lambda)$ is independent of $\lambda \in S$ and if for every $u \in D$, $\mathcal{L}(\lambda)u$ is holomorphic as a function of $\lambda \in S$.*

Note that a holomorphic pencil of densely-defined bounded operators (having by definition an operator-norm convergent power series expansion about each $\lambda_0 \in S$) is an example of a holomorphic family of type (A). In this context, we present the following notion of spectrum of operator pencils (see also [23, 40, 47]).

DEFINITION 2.4 (Spectrum of operator pencils and related notions, [57]). *Let $\mathcal{L}$ be a holomorphic family of type (A) on a Banach space X with domain D defined for $\lambda \in S \subset \mathbb{C}$. A complex number $\lambda_0 \in S$ is called a regular point of $\mathcal{L}$ if $\mathcal{L}(\lambda_0)$ has a bounded inverse on X. The set of all regular points of $\mathcal{L}$ is called the resolvent set $\rho(\mathcal{L})$ of $\mathcal{L}$. The complement of $\rho(\mathcal{L})$ in $S \subset \mathbb{C}$ is called the spectrum $\sigma(\mathcal{L})$ of $\mathcal{L}$. A complex number $\lambda_0 \in S$ is called a characteristic value of $\mathcal{L}$ if there exists a nonzero $u \in D$, called a characteristic vector corresponding to λ_0, satisfying*

$$\mathcal{L}(\lambda_0)u = 0. \tag{2.23}$$

The dimension of the kernel of $\mathcal{L}(\lambda_0)$ is called the geometric multiplicity of λ_0. A sequence of vectors $\{u^{[0]}, u^{[1]}, \dots, u^{[m-1]}\}$, where $u^{[0]}$ is a characteristic vector for λ_0, is called a chain of root vectors (or generalized characteristic vectors) of length m if

$$\sum_{k=0}^{j} \frac{1}{k!}\mathcal{L}^{(k)}(\lambda_0)u^{[j-k]} = 0,\ j = 1, \dots, m-1, \quad \mathcal{L}^{(k)}(\lambda_0)u := \frac{d^k}{d\lambda^k}(\mathcal{L}(\lambda)u)\bigg|_{\lambda=\lambda_0}. \tag{2.24}$$

A chain of length m that cannot be extended to a chain of length $m+1$ is called a maximal chain, and its vectors span a root space for the characteristic value λ_0. The sum of lengths of all maximal chains corresponding to a characteristic value λ_0 is called the algebraic multiplicity of λ_0. A characteristic value is called semi-simple if its geometric and algebraic multiplicities are finite and equal.

The characteristic values of $\mathcal{L}$ are contained in (but need not exhaust) $\sigma(\mathcal{L})$. Note that for holomorphic matrix pencils, the algebraic multiplicity of a characteristic value λ_0 reduces to the order of vanishing of the characteristic determinant $\det(\mathcal{L}(\lambda))$ at λ_0. However, in contrast to a Jordan chain of a non-semi-simple eigenvalue of a matrix, the root vectors making up a chain for an operator pencil need not be linearly independent, and a generalized characteristic vector can be identically equal to zero.

We can also consider the spectrum of $\mathcal{L}(\lambda)$ as an operator depending parametrically on $\lambda \in S$, and in particular the eigenvalues $\mu(\lambda)$ of $\mathcal{L}(\lambda)$. The fact that holomorphic families $\mathcal{L}(\lambda)$ of type (A) that have compact resolvent for some $\lambda_0 \in S$ have a compact resolvent for all $\lambda \in S$ [39, VII.2.1, p. 377, Theorem 2.4] can then be used to establish the following theorem.

THEOREM 2.5 (Kato, [39, VII.3.5, p. 392, Theorem 3.9]). *Let $\mathcal{L}(\lambda)$ be a selfadjoint holomorphic family of type (A) for $\lambda \in S$, and let $I_0 \subset S$ be a real interval. If $\mathcal{L}(\lambda_0)$ has a compact resolvent for some $\lambda_0 \in S$ then all the eigenvalues of $\mathcal{L}(\lambda)$ can be represented as a sequence $\{\mu_j(\lambda)\}$ of real holomorphic functions on I_0. Furthermore, the complete orthonormal family of corresponding eigenvectors can also be represented by a sequence of vector valued holomorphic functions on I_0.*

2.2.3. A special technique for polynomial operator pencils. Let $\mathcal{L}$ be a polynomial operator pencil of degree p: $\mathcal{L}(\lambda) := L_0 + \lambda L_1 + \cdots + \lambda^p L_p$, where L_j are all operators defined on a common dense domain D within a Banach space X, and suppose that L_p is invertible. Such a pencil is obviously a holomorphic family of type (A). In this case, the problem of characterizing the spectrum $\sigma(\mathcal{L})$ can be reduced to that of finding the spectrum (in the usual sense) of a single operator $C_{\mathcal{L}}$, called the *companion matrix* [23, 57], acting on the p-fold Cartesian product space X^p. Indeed, given $u \in D$, define $\mathbf{u} := (u, \lambda u, \ldots, \lambda^{p-1}u)^{\mathsf{T}} \in X^p$. Then the equation $\mathcal{L}(\lambda)u = 0$ on X is equivalent to the standard eigenvalue equation $C_{\mathcal{L}}\mathbf{u} = \lambda\mathbf{u}$ on X^p, where $C_{\mathcal{L}}$ is the $p \times p$ matrix of operators:

$$C_{\mathcal{L}} := \begin{pmatrix} 0 & \mathbb{I} & 0 & \cdots & 0 \\ 0 & 0 & \mathbb{I} & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & \mathbb{I} \\ -\tilde{L}_0 & -\tilde{L}_1 & -\tilde{L}_2 & \cdots & -\tilde{L}_{p-1} \end{pmatrix}, \quad \tilde{L}_k := L_p^{-1} L_k, \quad 0 \le k \le p-1. \tag{2.25}$$

Here $\mathbb{I}$ and 0 are respectively the identity and zero operators acting on X. The root spaces of the pencil $\mathcal{L}$ are the images of the invariant subspaces of the companion matrix $C_{\mathcal{L}}$ under the obvious projection of X^p onto the first factor X.

Since groups of isolated eigenvalues of $C_{\mathcal{L}}$ with finite multiplicities have properties similar to those of eigenvalues of finite-dimensional matrices, it is possible to completely characterize the chains of the pencil $\mathcal{L}$ in terms of the Jordan chains of $C_{\mathcal{L}}$. The constructive algebraic proof given for matrix pencils in [23] can be used without modification (see also [57, §12]).

THEOREM 2.6 ([23, p. 250, Proposition 12.4.1]). *Let $\mathcal{L}$ be a polynomial operator pencil of degree p with an invertible leading coefficient L_p, and let $C_{\mathcal{L}}$ be the companion*

matrix of $\mathcal{L}$. Then $\{u^{[0]},\dots,u^{[m-1]}\}$ is a chain of length m for $\mathcal{L}$ as defined in (2.24) at a characteristic value λ_0 if and only if the columns of the $p\times m$ matrix $V=(U,UJ_0,\dots,UJ_0^{p-1})^{\mathsf{T}}$ form a standard Jordan chain for $C_{\mathcal{L}}$ corresponding to the same λ_0, where J_0 is the $m\times m$ matrix Jordan block[7] with eigenvalue λ_0 and $U=(u^{[0]},\dots,u^{[m-1]})$.

2.3. Indefinite inner product spaces. Unfortunately, in the case that $\mathcal{L}$ is a *selfadjoint* polynomial operator pencil (with coefficients L_j being selfadjoint operators densely defined on a Hilbert space X), selfadjointness is lost in the extension process and the companion matrix $C_{\mathcal{L}}$ acting on X^p is non-selfadjoint with respect to the "Euclidean" inner product on X^p induced by that on X: $(\mathbf{u},\mathbf{v}):=(u_1,v_1)+\cdots+(u_p,v_p)$. However, a calculation shows that $C_{\mathcal{L}}$ is indeed selfadjoint with respect to an *indefinite* quadratic form defined by

$$\langle\mathbf{u},\mathbf{v}\rangle:=(\mathbf{u},B_{\mathcal{L}}\mathbf{v}),\quad B_{\mathcal{L}}:=\begin{pmatrix} L_1 & L_2 & \cdots & L_{p-1} & L_p\\ L_2 & L_3 & \cdots & L_p & 0\\ \vdots & \vdots & \iddots & \vdots & \vdots\\ L_{p-1} & L_p & \cdots & 0 & 0\\ L_p & 0 & \cdots & 0 & 0\end{pmatrix}. \tag{2.26}$$

Indeed, the Hankel-type operator matrix $B_{\mathcal{L}}$ (selfadjoint with respect to the Euclidean inner product on X^p) intertwines the companion matrix $C_{\mathcal{L}}$ with its adjoint $C_{\mathcal{L}}^*$ with respect to the Euclidean inner product as follows: $B_{\mathcal{L}}C_{\mathcal{L}}=C_{\mathcal{L}}^*B_{\mathcal{L}}$. This implies that the root spaces in X^p corresponding to different eigenvalues of $C_{\mathcal{L}}$ (characteristic values of $\mathcal{L}$) are orthogonal with respect to the quadratic form $\langle\cdot,\cdot\rangle$ defined by (2.26).

The extension technique therefore closely relates the theory of selfadjoint polynomial operator pencils to that of so-called *indefinite inner product spaces* or *Pontryagin spaces*. As is apparent from the definition (1.8), the Krein signature of a characteristic value is related to a certain indefinite quadratic form. Although one of the aims of our article is the avoidance of the algebra of indefinite inner product spaces, the latter are clearly lying just beneath the surface, so we would like to briefly cite some of the related literature. The seminal work of Pontryagin [66] opened up a huge field devoted to the spectral properties and invariant subspaces of operators in such spaces having a wide variety of applications. Important further developments of the theory were made in [32], and general references for many of the key results include [20, 23]. Over 40 years after its publication a central result of the theory — the Pontryagin Invariant Subspace Theorem — was rediscovered in connection with nonlinear waves and eigenvalue counting theorems [11, 25, 28] (see [11] for a historical discussion).

3. Graphical Interpretation of Multiplicity and Krein Signature. This section contains a survey of known results connecting the characteristic value problem (2.23) for a selfadjoint operator pencil $\mathcal{L}$ to the family, parametrized by $\lambda\in\mathbb{R}$, of selfadjoint eigenvalue problems

$$\mathcal{L}(\lambda)u(\lambda)=\mu(\lambda)u(\lambda). \tag{3.1}$$

3.1. Multiplicity of characteristic values. We first present a theorem well-known in the theory of matrix pencils ([23, §12.5, p. 259], see also [49, 57] for a similar approach in the case of operators). The key idea here is to relate the linear-algebraic

[7] $J_{0kl}:=\lambda_0\delta_{k-l}+\delta_{k-l+1}$ where δ_i denotes the Kronecker delta symbol.

notions of algebraic and geometric multiplicity of characteristic values and root vector chains to information contained in the graphs of eigenvector branches $\mu = \mu(\lambda)$ and the corresponding eigenvector branches $u = u(\lambda)$ of the selfadjoint eigenvalue problem (3.1). To give better insight into this result so fundamental to the approach described in this paper, we give its complete proof. A different proof can be found in [57, p. 57, Lemma 11.3, and p. 61, Corollary 12.4 and Lemma 12.5], where $\mu_i(\lambda)$ is called a generating polynomial (see also [43, 38]). The key property needed to establish the result is some degree of smoothness of the real functions $\mu = \mu(\lambda)$ at their roots; while it is not completely necessary, for simplicity we will assume that $\mathcal{L}$ is a selfadjoint operator pencil that (i) is a holomorphic family of type (A) and (ii) has compact resolvent for some (and hence for all) $\lambda_0 \in S$. Under these conditions we can appeal to Theorem 2.5 to guarantee the required smoothness. Our assumptions have the added benefit of making $\mathcal{L}(\lambda)$ a Fredholm operator for each $\lambda \in \mathbb{R} \subset S$. However, most of what will be formulated below will also hold if only $\mathcal{L}(\lambda)$ is Fredholm and the branches $\mu = \mu(\lambda)$ of interest are sufficiently smooth near a real characteristic value λ_0.

THEOREM 3.1 (Graphical interpretation of multiplicity). *Let $\mathcal{L}$ be a selfadjoint holomorphic family of type (A) with compact resolvent, and assume that $\mathcal{L}$ has an isolated real characteristic value λ_0. Then the following properties are equivalent:*

(a) λ_0 has finite algebraic multiplicity α and geometric multiplicity k with corresponding maximal chains of root vectors ($m_1 + \cdots + m_k = \alpha$)

$$\left\{u_1^{[0]}, \ldots, u_1^{[m_1-1]}\right\}, \left\{u_2^{[0]}, \ldots, u_2^{[m_2-1]}\right\}, \ \ldots, \ \left\{u_k^{[0]}, \ldots, u_k^{[m_k-1]}\right\}. \quad (3.2)$$

(b) There exist exactly k analytic eigenvalue branches $\mu = \mu_1(\lambda), \ldots, \mu = \mu_k(\lambda)$, vanishing at $\lambda = \lambda_0$ where $\mu_j(\lambda)$ vanishes to order m_j: $\mu_j^{(n)}(\lambda_0) = 0$ for $n = 0, 1, \ldots, m_j - 1$, while $\mu_j^{(m_j)}(\lambda_0) \neq 0$. The derivatives of the corresponding orthonormal analytic eigenvector branches $u = u_1(\lambda), \ldots, u = u_k(\lambda)$ encode the chains of root vectors as follows:

$$u_j^{[r]} = \frac{1}{r!}\frac{d^r u_j}{d\lambda^r}(\lambda_0), \quad r = 0, 1, \ldots, m_j - 1. \quad (3.3)$$

Proof. Note that $\mathcal{L}(\lambda_0)$ is a Fredholm operator, so the Fredholm alternative applies, and that analyticity of all eigenvalue and eigenvector branches is guaranteed by Theorem 2.5. (The theorem holds true under more general conditions that guarantee these properties of $\mathcal{L}$.) We prove the theorem by mathematical induction on $m = \max_{1 \le j \le k} m_j$.

Let $m = 1$ (and hence $m_j = 1$ for $j = 1, \ldots, k$). First suppose that (b) holds, which implies that $\operatorname{Ker} \mathcal{L}(\lambda_0)$ is spanned by the vectors $u_1(\lambda_0), \ldots, u_k(\lambda_0)$ and one can set $u_j^{[0]} := u_j(\lambda_0)$, $j = 1, \ldots, k$. It remains to show that each of the single-vector chains $\{u_j^{[0]}\}$ is maximal. Our assumptions allow us to differentiate (3.1) with respect to λ along the branches $\mu = \mu_j(\lambda)$ and $u = u_j(\lambda)$ and evaluate at $\lambda = \lambda_0$ to obtain

$$\mathcal{L}'(\lambda_0)u_j^{[0]} + \mathcal{L}(\lambda_0)u_j'(\lambda_0) = \mu_j'(\lambda_0)u_j^{[0]} \quad (3.4)$$

where we have used $\mu_j(\lambda_0) = 0$. Note that as $m_j = 1$ we have $\mu_j'(\lambda_0) \neq 0$. If the chain $\{u_j^{[0]}\}$ were not maximal, then by Definition 2.4 there would exist a vector $u_j^{[1]}$ such that $\mathcal{L}(\lambda_0)u_j^{[1]} + \mathcal{L}'(\lambda_0)u_j^{[0]} = 0$. Since $\mathcal{L}(\lambda_0)u_j^{[0]} = 0$ and $(u_j^{[0]}, u_j^{[0]}) = 1$, we can

combine this with (3.4) to obtain

$$0 = \left(u_j'(\lambda_0) - u_j^{[1]}, \mathcal{L}(\lambda_0)u_j^{[0]}\right) = \left(\mathcal{L}(\lambda_0)\left[u_j'(\lambda_0) - u_j^{[1]}\right], u_j^{[0]}\right) = \mu_j'(\lambda_0) \neq 0, \quad (3.5)$$

an obvious contradiction. Therefore each chain $\{u_j^{[0]}\}$ for $j = 1, \ldots, k$, is maximal, and we conclude that (a) holds.

On the other hand, if (a) holds, then $\operatorname{Ker}\mathcal{L}(\lambda_0)$ has dimension k, and hence there exist k eigenvalue branches $\mu = \mu_j(\lambda)$ (counted with multiplicity), all analytic functions with roots at $\lambda = \lambda_0$. To conclude that (b) holds it just remains to prove that each $\mu_j(\lambda)$ vanishes at $\lambda = \lambda_0$ to exactly first order. But were this not the case for some j, we would have $\mu_j'(\lambda_0) = 0$ and then (3.4) would imply the existence of a generalized root vector $u_j^{[1]} := u_j'(\lambda_0)$, in contradiction with the fact guaranteed by case (a) with $m_j = m = 1$ that λ_0 is semi-simple. The theorem therefore holds in the case that $m = 1$.

Now suppose that the theorem is true for some $m = M \geq 1$. Let $m = M + 1$ and assume that (b) holds. Let $\mu = \mu_j(\lambda)$ be an eigenvalue branch that vanishes at $\lambda = \lambda_0$ to order $M + 1$. We differentiate (3.1) M times with respect to λ along $\mu = \mu_j(\lambda)$ and $u = u_j(\lambda)$ at $\lambda = \lambda_0$ to obtain

$$\sum_{n=0}^{M} \binom{M}{n} \mathcal{L}^{(n)}(\lambda_0) u_j^{(M-n)}(\lambda_0) = 0, \quad (3.6)$$

since $\mu_j^{(n)}(\lambda_0) = 0$ for $n = 0, \ldots, M$. By the inductive hypothesis, a chain of root vectors of length M for the characteristic value λ_0 is given by

$$\left\{u_j^{[0]}, u_j^{[1]}, \ldots, u_j^{[M-1]}\right\} = \left\{u_j(\lambda_0), u_j'(\lambda_0), \ldots, \frac{1}{(M-1)!} u_j^{(M-1)}(\lambda_0)\right\}. \quad (3.7)$$

Using (3.7) in (3.6) and comparing with Definition 2.4 shows that the chain can be extended to one of length $M + 1$ by defining $u_j^{[M]} := u_j^{(M)}(\lambda_0)/M!$. That the extended chain is maximal follows from the fact that $\mu_j^{(M+1)}(\lambda_0) \neq 0$ just as in the case of $m = 1$, and therefore (a) holds for $m = M + 1$.

Now assume instead that (a) holds for $m = M + 1$ and that the theorem is true for $m = M$. Then for each j with $m_j = M + 1$ there exists a generalized root vector $u_j^{[M]}$ satisfying (according to Definition 2.4 and the inductive hypothesis)

$$\sum_{n=0}^{M} \frac{1}{n!} \mathcal{L}^{(n)}(\lambda_0) u_j^{[M-n]} = \mathcal{L}(\lambda_0) u_j^{[M]} + \frac{1}{M!} \sum_{n=1}^{M} \binom{M}{n} \mathcal{L}^{(n)}(\lambda_0) u_j^{(M-n)}(\lambda_0) = 0. \quad (3.8)$$

The inductive hypothesis implies that $\mu_j^{(n)}(\lambda_0) = 0$ for $n = 0, 1, \ldots, M-1$, so in place of (3.6) we have

$$\sum_{n=0}^{M} \binom{M}{n} \mathcal{L}^{(n)}(\lambda_0) u_j^{(M-n)}(\lambda_0) = \mu_j^{(M)}(\lambda_0) u_j^{[0]}. \quad (3.9)$$

Comparing these two equations shows that

$$\mathcal{L}(\lambda_0)\left(u_j^{(M)}(\lambda_0) - M! u_j^{[M]}\right) = \mu_j^{(M)}(\lambda_0) u_j^{[0]}. \quad (3.10)$$

Thus $\mu_j^{(M)}(\lambda_0)u_j^{[0]} \in \mathrm{Ker}(\mathcal{L}(\lambda_0)) \cap \mathrm{Ran}(\mathcal{L}(\lambda_0))$ and by $\mathcal{L}(\lambda_0)^* = \mathcal{L}(\lambda_0)$ for $\lambda_0 \in \mathbb{R}$ the Fredholm alternative implies[8] that $\mu_j^{(M)}(\lambda_0) = 0$ because $u_j^{[0]} \neq 0$. The proof of non-vanishing of $\mu_j^{(M+1)}(\lambda_0)$ using the maximality of the chain is similar to the case of $m = 1$. Therefore (b) holds as well.

By mathematical induction it therefore follows that the theorem is true for all finite $m = \max_{1\leq j\leq k} m_j$. □

The significance of Theorem 3.1 is that the algebraic and geometric multiplicities of isolated real characteristic values of selfadjoint pencils can be simply read off from plots of the real eigenvalue branches $\mu = \mu_j(\lambda)$ of the selfadjoint eigenvalue problem (3.1). Indeed, real characteristic values are simply the locations of the intercepts of the branches with $\mu = 0$, the geometric multiplicity of a real characteristic value λ_0 is simply the number of branches crossing at the point $(\lambda, \mu) = (\lambda_0, 0)$, each branch $\mu = \mu_j(\lambda)$ corresponds to a single maximal chain of generalized root vectors, and the length of each maximal chain is simply the order of vanishing of the analytic function $\mu_j(\lambda)$ at $\lambda = \lambda_0$.

3.2. Krein signature of real characteristic values. The Krein signature of a real characteristic value is usually defined in terms of the restriction of an appropriate indefinite quadratic form to a root space, with the signature being positive, negative, or indefinite according to the type of the restricted form. Our goal is to give a useful definition of Krein signatures for real characteristic values of quite general operator pencils, however at the beginning we will restrict ourselves to polynomial pencils with invertible leading coefficient, as this makes available the companion matrix method described in § 2.2.3. In this case the relevant quantities can be defined as follows.

Definition 3.2 (Krein indices and Krein signature). *Let $\mathcal{L}$ be a selfadjoint polynomial operator pencil of degree p with invertible leading coefficient L_p acting in a Hilbert space X, and let λ_0 be an isolated real characteristic value of $\mathcal{L}$. Given a root space $\mathcal{U} \subset X$ spanned by the vectors of a maximal chain $\{u^{[0]}, \ldots, u^{[m-1]}\}$ for λ_0, let $\mathbf{u}^{[j]} \in X^p$, $j = 0, \ldots, m-1$, denote the columns of the matrix V defined in the statement of Theorem 2.6, and let W be the $m \times m$ Gram matrix with elements $W_{jk} := \langle \mathbf{u}^{[j]}, \mathbf{u}^{[k]} \rangle$, in terms of the indefinite Hermitian quadratic form $\langle \cdot, \cdot \rangle$ given by (2.26). The number of positive (negative) eigenvalues of the Hermitian matrix W is called the positive (negative) Krein index of the root space[9] $\mathcal{U}$ at λ_0 and is denoted $\kappa^+(\mathcal{U}, \lambda_0)$ ($\kappa^-(\mathcal{U}, \lambda_0)$). The sums of $\kappa^\pm(\mathcal{U}, \lambda_0)$ over all root spaces $\mathcal{U}$ of λ_0 are called the positive and negative Krein indices of λ_0 and are denoted $\kappa^\pm(\lambda_0)$. Finally, $\kappa(\mathcal{U}, \lambda_0) := \kappa^+(\mathcal{U}, \lambda_0) - \kappa^-(\mathcal{U}, \lambda_0)$ is called the Krein signature of the root space $\mathcal{U}$ for λ_0, and $\kappa(\lambda_0) := \kappa^+(\lambda_0) - \kappa^-(\lambda_0)$ is called the Krein signature of λ_0.*

If $\kappa^+(\lambda_0) > 0$ (respectively $\kappa^-(\lambda_0) > 0$) we say that λ_0 has a non-trivial positive (respectively negative) signature, and if both $\kappa^+(\lambda_0)$ and $\kappa^-(\lambda_0)$ are positive we say that λ_0 has indefinite Krein signature. It is not obvious but true[10] that the Krein signature of a single root space $\kappa(\mathcal{U}, \lambda_0)$ is always either 1, -1, or 0 regardless of the dimension of $\mathcal{U}$. The "total" Krein signature $\kappa(\lambda_0)$ of a real characteristic value λ_0 can, however, take any integer value by this definition. If λ_0 is a simple real characteristic value with characteristic vector $u \in X$, then the only root space $\mathcal{U}$ is spanned by u, and

[8] If $\lambda_0 \notin \mathbb{R}$ the argument fails, as the Fredholm alternative only implies that $\mu_j^{(M)}(\lambda_0)u_j^{[0]} \in \mathrm{Ker}(\mathcal{L}(\lambda_0))$ and $\mu_j^{(M)}(\lambda_0)u_j^{[0]} \perp \mathrm{Ker}(\mathcal{L}(\overline{\lambda}_0))$. Such a nontrivial vector can indeed exist.

[9] The Krein indices $\kappa^\pm(\mathcal{U}, \lambda_0)$ are in fact well-defined although the construction appears to depend on the specific choice of vectors making up the chain that spans the space $\mathcal{U}$.

[10] It is part of the proof of Theorem 3.4 below. See [43].

the corresponding (only) column of the matrix V is $\mathbf{u} = (u, \lambda_0 u, \dots, \lambda_0^{p-1} u)^\mathsf{T} \in X^p$, so the corresponding Gram matrix W is a real scalar given by

$$W = \langle \mathbf{u}, \mathbf{u} \rangle = (L_1 u + 2\lambda_0 L_2 u + \cdots + p\lambda_0^{p-1} L_p u, u) = (\mathcal{L}'(\lambda_0) u, u) \tag{3.11}$$

as a simple calculation using (2.26) shows. If furthermore $\mathcal{L}$ is a linear selfadjoint pencil, $\mathcal{L}(\lambda) = \lambda L_1 + L_0$, then $\mathcal{L}'(\lambda_0) = L_1$ and the Krein signature is given by $\kappa(\lambda_0) = \operatorname{sign}(L_1 u, u)$, which coincides with the usual definition (see (1.8)) found frequently in the literature.

This definition mirrors a recent approach [36, 38, 54] to spectral analysis of infinite-dimensional linearized Hamiltonian systems in which the important information about a point of the spectrum is obtained from a finite-dimensional reduction called a *Krein matrix*, an analogue of the Gram matrix W. Kollár and Pego [44] developed a rigorous perturbation theory of Krein signatures for finite systems of eigenvalues in the spirit of Kato [39], proving results that had previously belonged to the mathematical folklore in the field.

3.3. Graphical Krein signature theory. One of the main messages of our paper is that the correct way to generalize the notion of Krein indices and signatures so that they apply to and are useful in the analysis of spectral problems involving operator pencils of non-polynomial type like that illustrated in Example 8 is to eschew linear algebra in favor of analysis of eigenvalue curves $\mu = \mu(\lambda)$ of the problem (3.1) in the vicinity of their roots. This approach is attractive even in cases where companion matrix methods explained in §2.2.3 suffice to define the relevant quantities. We therefore begin by formulating such a "graphical" definition of Krein signatures.

DEFINITION 3.3 (Graphical Krein indices and signatures). *Let $\mathcal{L}(\lambda)$ be a self-adjoint holomorphic family of type (A) with compact resolvent, and assume that $\mathcal{L}$ has an isolated real characteristic value λ_0. Let $\mu = \mu(\lambda)$ be one of the real analytic eigenvalue branches of the problem (3.1) with $\mu^{(n)}(\lambda_0) = 0$ for $n = 0, 1, \dots, m-1$, while $\mu^{(m)}(\lambda_0) \neq 0$. Let $\eta(\mu) := \operatorname{sign}(\mu^{(m)}(\lambda_0)) = \pm 1$. Then the quantities*

$$\kappa_{\mathrm{g}}^{\pm}(\mu, \lambda_0) := \begin{cases} \frac{1}{2} m, & \text{for } m \text{ even} \\ \frac{1}{2}(m \pm \eta(\mu)), & \text{for } m \text{ odd} \end{cases} \tag{3.12}$$

are called the positive and negative graphical Krein indices of the eigenvalue branch $\mu = \mu(\lambda)$ corresponding to the characteristic value λ_0. The sums of $\kappa_{\mathrm{g}}^{\pm}(\mu, \lambda_0)$ over all eigenvalue branches crossing at $(\lambda, \mu) = (\lambda_0, 0)$ are called the positive and negative graphical Krein indices of λ_0 and are denoted $\kappa_{\mathrm{g}}^{\pm}(\lambda_0)$. Finally, $\kappa_{\mathrm{g}}(\mu, \lambda_0) := \kappa_{\mathrm{g}}^{+}(\mu, \lambda_0) - \kappa_{\mathrm{g}}^{-}(\mu, \lambda_0)$ is called the graphical Krein signature of the eigenvalue branch $\mu = \mu(\lambda)$ vanishing at λ_0, and $\kappa_{\mathrm{g}}(\lambda_0) := \kappa_{\mathrm{g}}^{+}(\lambda_0) - \kappa_{\mathrm{g}}^{-}(\lambda_0)$ is called the graphical Krein signature of λ_0.

Note that it follows directly from the definition of the graphical Krein indices $\kappa_{\mathrm{g}}^{\pm}(\mu, \lambda_0)$ that the graphical Krein signature $\kappa_{\mathrm{g}}(\mu, \lambda)$ necessarily takes one of the three values 1, −1, or 0. The concept of *sign characteristics* of operator pencils [23, Chapters 7 and 12] is closely related to our definition; here each root space of $\mathcal{L}$ corresponding to a characteristic value λ_0 is associated with the sign of $\tilde{\mu}(\lambda_0) \neq 0$, where $\mu(\lambda) = (\lambda - \lambda_0)^{m-1} \tilde{\mu}(\lambda)$ near $\lambda = \lambda_0$. See also [23, Theorem 5.11.2]. The relationship between sign characteristics and Krein indices can be expressed by an algebraic formula [23, p. 77, equation (5.2.4)].

Due to this correspondence, the Krein signature as given by Definition 3.2 and the graphical Krein signature as given by Definition 3.3 are well-known to be connected

in the theory of matrix pencils (see (1.10) for the relevant calculation in the case of a simple characteristic value). The connection is described in detail in [23, p. 260, Theorem 12.5.2] (see also [21, §10.5]). The proof as formulated in these works for matrix pencils is essentially the same as for operator pencils. A different proof of a more perturbative nature can be found in [36, Lemma 9]. See also [11, Lemma 5.3] and [72, Theorem 1] for the same result in the special case a generalized eigenvalue problem (linear pencil). Yet another proof based on the Frobenius rule can be found in [43, Theorem 4.2].

THEOREM 3.4 (Graphical nature of Krein indices and signatures). *Let $\mathcal{L}$ be a self-adjoint polynomial operator pencil of degree p with invertible leading coefficient L_p, and assume also that $\mathcal{L}$ is a holomorphic family of type (A) with compact resolvent. Let λ_0 be an isolated real characteristic value of $\mathcal{L}$. Let $\mathcal{U} = \operatorname{span}\{u^{[0]}, u^{[1]}, \dots, u^{[m-1]}\}$ be a the root space for λ_0 corresponding (according to Theorem 3.1) to the analytic eigenvalue branch $\mu = \mu(\lambda)$ of the associated eigenvalue problem* (3.1). *Then*

$$\kappa^{\pm}(\mathcal{U}, \lambda_0) = \kappa_{\mathrm{g}}^{\pm}(\mu, \lambda_0), \tag{3.13}$$

from which it follows that

$$\kappa(\mathcal{U}, \lambda_0) = \kappa_{\mathrm{g}}(\mu, \lambda_0) \tag{3.14}$$

and summing over all root spaces $\mathcal{U}$ and corresponding eigenvalue branches μ,

$$\kappa^{\pm}(\lambda_0) = \kappa_{\mathrm{g}}^{\pm}(\lambda_0) \quad \textit{and} \quad \kappa(\lambda_0) = \kappa_{\mathrm{g}}(\lambda_0). \tag{3.15}$$

In particular, this result implies that for simple characteristic values λ_0, the Krein signature can be calculated easily from the formula (1.9). This result allows us to extend the notion of Krein signature in a very natural way to problems arising in the theory of stability of nonlinear waves that cannot easily be formulated as spectral problems for polynomial pencils as the spectral parameter λ enters into the problem in two different, transcendentally related, ways (as in Example 8). Finally, the graphical Krein signature is the one that is most easily detected by a natural generalization of the Evans function, the subject that we take up next in §4. From now on, when we refer to total Krein indices and Krein signatures of a real characteristic value λ_0, we will always mean the (more widely applicable) graphical quantities of Definition 3.3, and omit the unnecessary subscript "g".

4. Evans Functions. Consider a dynamical system linearized about an equilibrium in the form (compare to (1.2))

$$\frac{du}{dt} = Au. \tag{4.1}$$

In the infinite-dimensional context typical in nonlinear wave theory, the linearized system (4.1) is usually a partial differential equation and u is a vector in some Hilbert space X. For waves in $1+1$ dimensions, X is a space of functions of a real spatial variable x, and A can be thought of as a linear differential operator acting in X. The key spectral problem in the stability analysis of the equilibrium is (compare to (1.1))

$$-i\lambda u_0 = Au_0, \tag{4.2}$$

as each eigenvector $u_0 = u_0(\cdot) \in X$ of (4.2) generates a solution of (4.1) in the separated form $u = u(x,t) = e^{-i\lambda t}u_0(x)$. Values of $\nu = -i\lambda \in \sigma(A)$ having a positive

real part imply the existence of exponentially growing modes in (4.1) and thus linear instability of the equilibrium of the original (nonlinear) dynamical system.

Although numerical methods to study $\sigma(A)$ based on spatial discretization or some other finite dimensional truncation are easy to implement, they may fail to detect the full extent of the unstable spectrum or they may introduce spurious λ, particularly in the vicinity of any continuous spectrum of A. In [15]–[18] a more robust and reliable numerical method was developed based on a new concept now called the *Evans function*, and this method was successfully applied to study the stability of neural impulses. Since its introduction, the Evans function has become a popular tool for the detection of stability of various types of waves in many applications including fluid mechanics, condensed matter physics, combustion theory, etc.

4.1. Typical construction and use of Evans functions. We begin with a standard definition.

DEFINITION 4.1 (Evans functions). *Let $\mathcal{L} = \mathcal{L}(\lambda)$ be an operator pencil. An analytic function $D : \Omega \subset \mathbb{C} \to \mathbb{C}$ whose roots in Ω coincide exactly with isolated characteristic values $\lambda = \lambda_0$ of the spectral problem (2.23), and that vanishes at each such point to precisely the order of the algebraic multiplicity of the characteristic value is called an Evans function for $\mathcal{L}$ (on Ω).*

It is common to refer to "the" Evans function for a given spectral problem, and this usually implies a particular kind of construction appropriate for problems of the special form (4.2) that we will describe briefly below. However the key properties of analyticity and vanishing on the discrete spectrum with the correct multiplicity are shared by many other functions and as the freedom to choose among them can be useful in applications, we prefer to keep the terminology as general as possible.

Let us now describe the classical construction of "the" Evans function for (4.2) in the case that A is a scalar ordinary differential operator of order k acting in X. To begin with, the spectral problem (4.2) is rewritten as a (nonautonomous, in the usual case that A has non-constant coefficients) first-order system[11]

$$dv/dx = B(x,\lambda)v, \qquad v \in X^k\,, \tag{4.3}$$

where $B(x,\lambda)$ is a $k \times k$ matrix-valued function assumed to take finite limits $B_\pm(\lambda)$ as $x \to \pm\infty$. We consider (4.3) along with the constant-coefficient "asymptotic systems"

$$dv/dx = B_\pm(\lambda)v. \tag{4.4}$$

The first-order system (4.3) has a nonzero solution $v(x)$ decaying as $|x| \to \infty$, and hence (4.2) has a nontrivial solution $u \in X$, exactly for those values of λ for which the forward evolution $W_{\rm u}^{-\infty}(x)$ of the unstable manifold of the zero equilibrium of (4.4) for $B_-(\lambda)$ has a transversal intersection with the backward evolution $W_{\rm s}^{+\infty}(x)$ of the stable manifold of the zero equilibrium of (4.4) for $B_+(\lambda)$. To properly define the evolutes $W_{\rm u}^{-\infty}(x)$ and $W_{\rm s}^{+\infty}(x)$ for a common value of $x \in \mathbb{R}$ requires that both asymptotic systems (4.4) are hyperbolic with "exponential dichotomy" (that is, the eigenvalues of $B_\pm(\lambda)$ are bounded away from the imaginary axis as λ varies in the region of interest), and for the existence of isolated characteristic values λ one usually requires complementarity (in $\mathbb{C}^k$) of the dimensions of the stable and

[11] There are of course many ways to rewrite a single higher-order linear differential equation as a first-order system, and if care is not taken key symmetries of the original equation can be lost in the process. This is true even in the case of (1.1), although for such problems with Hamiltonian symmetry Bridges [6] has shown how some of this structure can be retained.

the unstable manifolds at $x = \pm\infty$. The traditional Evans function $D(\lambda)$ detecting transversality of the intersection of the evolutes can then be expressed as a $k \times k$ Wronskian determinant whose columns include $p_- < k$ vectors spanning $W_{\rm u}^{-\infty}(x)$ and $p_+ := k - p_-$ vectors spanning $W_{\rm s}^{+\infty}(x)$. For coefficient matrices $B(x,\lambda)$ of trace zero, Abel's Theorem implies that $D(\lambda)$ is independent of the value of $x \in \mathbb{R}$ at which the Wronskian is computed. For systems with nonzero but smooth trace, *any* value of $x \in \mathbb{R}$ can be chosen but different values of x lead to different Evans functions $D(\lambda)$ the ratio of any two of which is an analytic non-vanishing function of λ. Usually in such situations the unimportant dependence of $D(\lambda)$ on $x \in \mathbb{R}$ is explicitly removed by an appropriate exponential normalization [63]. This type of construction has been given a topological interpretation by Alexander *et al.* [1], who related $D(\lambda)$ to the Chern number of a tangential fibre bundle. Pego and Weinstein [64] used this definition of $D(\lambda)$ to establish stability properties of solitary waves in nonlinear systems including the Korteweg-deVries, Benjamin-Bona-Mahoney, and Boussinesq equations and also pointed out a connection between $D(\lambda)$ and the *transmission coefficient* associated with *Jost solutions* in scattering theory (see also [75]). This type of construction has also been extended to problems in multiple dimensions where A is a partial differential operator; in these cases special steps must be taken to ensure the analyticity of the resulting Evans function. In low dimensions a useful generalization of the Wronskian can be constructed using exterior products [63], and a robust numerical algorithm has also been developed [31] for this purpose that uses continuous orthogonalization to obtain an analytic Evans function.

The main reason for insisting upon the key property of analyticity of an Evans function is that the presence of characteristic values, as zeros of the analytic function $D(\lambda)$, can be detected *from a distance* in the complex plane by means of the Argument Principle. Indeed, to determine the number of characteristic values (counting algebraic multiplicity) in a two-dimensional simply-connected bounded domain Ω of the complex plane, it suffices to evaluate an Evans function only on the one-dimensional boundary $\partial\Omega$ and to compute the winding number of the phase of $D(\lambda)$ as λ traverses this curve. If the winding number is zero, one knows that there are no characteristic values in Ω. Otherwise, Ω can be split into subregions and the process repeated to further confine the discrete spectrum. From the computational point of view this is both more robust and far less expensive than a fully two-dimensional search of the λ-plane for characteristic values. Searching for characteristic values by computing the winding number of an Evans function $D(\lambda)$ along closed curves is a particularly useful approach to the spectral problem (2.23) if one knows, say by a variational argument, a bound of the form $|\lambda| < M$ for the characteristic values of the problem. However, even in cases when such a bound in unavailable, it turns out that if the Krein signature of certain characteristic values (usually the real ones) is known or can be computed it is sometimes possible to justify numerics and to reduce computational costs significantly [44]. This suggests that it would be particularly useful if it were possible to extract the Krein signature of a real characteristic value directly from an Evans function.

4.2. Evans functions and Krein signatures. By definition, an Evans function $D(\lambda)$ detects characteristic values of (2.23) but the usual constructions produce Evans functions that do not provide or contain any information about the corresponding root vectors. To better understand the problem, first consider a simple finite-dimensional case of the linearized Hamiltonian spectral problem (1.1) and the related spectral problem (1.4) for an equivalent linear pencil $\mathcal{L}(\lambda) = L - \lambda K$. In this

case the most obvious definition of an Evans function is to set

$$D(\lambda) := \det(\mathcal{L}(\lambda)) = \det(L - \lambda K). \tag{4.5}$$

Note that as L and K are Hermitian matrices, D is real-valued as a function of $\lambda \in \mathbb{R}$ but of course it has meaning for complex λ as well. It might seem natural to expect that the Evans function also encodes the Krein signature of the characteristic values in some sense. It has frequently been thought that under some circumstances at least the sign of the slope $D'(\lambda_0)$ may be correlated with the Krein signature $\kappa(\lambda_0)$. To explore this idea, we present the following two examples illustrating the graphical interpretation of the Krein signature as explained in Definition 3.3 and Theorem 3.4, and comparing with the slope of $D(\lambda)$ at its roots.

◁ *Example 9.* Let L and J be given by (compare to Example 2)

$$L = \begin{pmatrix} 2 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -2 \end{pmatrix}, \quad J = \begin{pmatrix} 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 1 \\ -2 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{pmatrix}.$$

The four real eigenvalue branches $\mu = \mu(\lambda)$ of the associated spectral problem (3.1) for this example are plotted along with the graph of $D(\lambda)$ as given by (4.5) in the left-hand panel of Fig. 4.1. The Krein signature of each of the real characteristic values

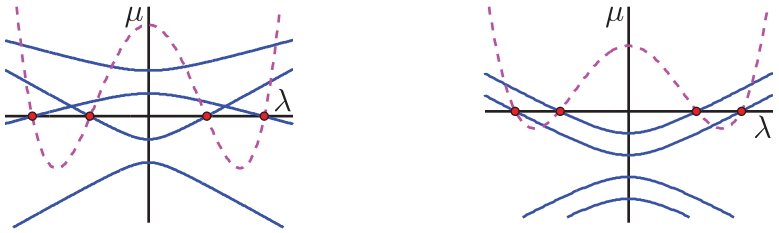

FIG. 4.1. *The four eigenvalue branches $\mu = \mu(\lambda)$ of the linear pencil $\mathcal{L}(\lambda)$ for Examples 9–10 are plotted with solid curves. The real intercepts of the branches with $\mu = 0$ corresponding to real characteristic values of $\mathcal{L}$ are indicated with dots. The dashed curve is the graph of the Evans function (arbitrary units) $D(\lambda) = \det(\mathcal{L}(\lambda))$.*

$\lambda = \lambda_0$ can be easily calculated graphically using Definition 3.3 and Theorem 3.4, or equivalently as all four characteristic values are simple, via (1.9). In this example it is easy to see that for each of the four real characteristic values, $\kappa(\lambda_0) = -\text{sign}(D'(\lambda_0))$, and therefore the sign of the slope $D'(\lambda)$ of $D(\lambda)$ at its zeros is strongly correlated with the Krein signature. ▷

◁ *Example 10.* Let L and J be given by (compare to Example 1)

$$L = \begin{pmatrix} -0.5 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1.5 & 0 \\ 0 & 0 & 0 & -2 \end{pmatrix}, \quad J = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ -1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{pmatrix}.$$

This case is illustrated in the right-hand panel of Fig. 4.1. There are again four simple real characteristic values of $\mathcal{L}$ (here consistent with the fact that L is negative

definite), and by using the graphical definition (1.9) we see that the positive (negative) characteristic values each have Krein signature $\kappa(\lambda_0) = +1$ ($\kappa(\lambda_0) = -1$). By contrast with Example 9, in this case it is obvious that the slope of $D(\lambda)$ at its roots is positively correlated with the Krein signature of half of the characteristic values and is negatively correlated with the Krein signature of the other half. Hence $D'(\lambda)$ does not detect the Krein signature correctly as the Evans function fails to capture the attachment of the characteristic values to different eigenvalue branches $\mu = \mu(\lambda)$. $\triangleright$

These simple examples illustrate the state of understanding in much of the literature: it just is not clear under which circumstances one may expect derivatives of an Evans function to encode Krein signatures of characteristic values.

4.3. Evans-Krein functions. In the general setting of operator pencils we now address the question posed in [37] (in the context of linearized Hamiltonians) as to whether it is possible to determine Krein signatures from an Evans function. In fact we can answer this question in the affirmative by showing how a simple generalization of the notion of an Evans function correctly detects the Krein signature in every case. Our approach provides a simple way to calculate Krein signatures of real characteristic values of selfadjoint operator pencils by an elementary augmentation of virtually any existing Evans function evaluation code at almost no additional programming cost. The main idea is to build into an Evans function information about the individual eigenvalue branches $\mu = \mu(\lambda)$ to which the real characteristic values are associated as roots. In fact, this simply amounts to bringing in μ as an additional parameter. We now introduce such an appropriate generalization of an Evans function, which we call an *Evans-Krein function*.

DEFINITION 4.2 (Evans-Krein functions). *Let $\mathcal{L} = \mathcal{L}(\lambda)$ be an operator pencil acting on a space X, and consider the related pencil $\mathcal{M} = \mathcal{M}(\lambda;\mu) := \mathcal{L}(\lambda) - \mu I$ where I denotes the identity operator on X and where μ is a sufficiently small real parameter. An Evans function $E(\lambda;\mu)$ (in the sense of Definition 4.1) for the μ-parametrized pencil $\mathcal{M}$ is called an Evans-Krein function for $\mathcal{L}$.*

Under suitable general conditions on $\mathcal{L}$, an Evans-Krein function is not only analytic for $\lambda \in \Omega \subset \mathbb{C}$ for each small μ, but also will be analytic in both λ and μ on a set of the form $\Omega \times \{|\mu| < \delta\} \subset \mathbb{C}^2$ for some $\delta > 0$. Upon setting $\mu = 0$ we see that if $E(\lambda;\mu)$ is an Evans-Krein function for $\mathcal{L}$, then $D(\lambda) := E(\lambda;0)$ is an Evans function for $\mathcal{L}$, with all of the coincident properties as given in Definition 4.1. It is also clear that if $\mu = \mu(\lambda)$ is a real eigenvalue branch of the problem (3.1), then $E(\lambda;\mu(\lambda))$ vanishes identically for all $\lambda \in \mathbb{C}$ for which the branch is defined and $\mu(\lambda)$ is sufficiently small. Moreover, if $\lambda_0 \in \mathbb{R}$ is a real characteristic value of $\mathcal{L}$ of geometric multiplicity $k \geq 1$, and if $\mu_1(\lambda), \ldots, \mu_k(\lambda)$ are the corresponding eigenvalue branches of (3.1) crossing at the point $(\lambda, \mu) = (\lambda_0, 0)$, then in a $\mathbb{C}^2$ neighborhood of this point, an arbitrary Evans-Krein function can be locally represented in the normal form

$$E(\lambda;\mu) = C(\lambda;\mu) \prod_{j=1}^{k} (\mu - \mu_j(\lambda)) \tag{4.6}$$

where $C(\lambda;\mu)$ is an analytic function with $C(\lambda_0;0) \neq 0$.

Construction of an Evans-Krein function in practice can be very similar to that of an Evans function in the usual sense. For a matrix pencil $\mathcal{L}$ one can set

$$E(\lambda;\mu) := \det(\mathcal{L}(\lambda) - \mu\mathbb{I}), \quad \text{and} \quad D(\lambda) := E(\lambda;0). \tag{4.7}$$

This definition is also suitable if $\mathcal{L}$ is an operator pencil that is a trace-class perturbation of the identity, in which case the determinants above are to be understood as Fredholm determinants. In the case that the spectral problem (2.23) involves differential operators and one is led to consider the traditional construction of an Evans function (as described in §4.1) via a Wronskian of subspaces of solutions decaying at the opposite spatial infinities, *exactly the same construction gives an Evans-Krein function if the pencil $\mathcal{L}$ is merely perturbed into $\mathcal{M}$*, because for sufficiently small μ it is easy to guarantee the persistence of hyperbolicity of the asymptotic systems (4.4) as well as the complementarity of the stable and unstable subspace dimensions.

4.4. Calculation of Krein signatures from an Evans-Krein function. Let $E(\lambda;\mu)$ be an Evans-Krein function for a selfadjoint operator pencil $\mathcal{L}$ for which eigenvalue branches $\mu = \mu(\lambda)$ of the associated problem (3.1) are sufficiently smooth, and let $\lambda = \lambda_0$ be an isolated real characteristic value of $\mathcal{L}$. By differentiating the identity $E(\lambda;\mu(\lambda)) = 0$ with respect to λ along a particular smooth branch $\mu = \mu(\lambda)$ for which $\mu(\lambda_0) = 0$ and then evaluating at $\lambda = \lambda_0$ we easily obtain

$$E_\lambda(\lambda_0;0) + E_\mu(\lambda_0;0)\mu'(\lambda_0) = 0. \tag{4.8}$$

According to the normal form (4.6), the partial derivatives $E_\lambda(\lambda_0;0)$ and $E_\mu(\lambda_0;0)$ are both nonzero under the assumption that the characteristic value λ_0 is simple, which implies both that the number of eigenvalue branches $\mu = \mu(\lambda)$ of (3.1) crossing the $\mu = 0$ axis at $\lambda = \lambda_0$ is exactly $k = 1$, and also that λ_0 is a simple root of the branch function $\mu(\lambda)$ (see §3 for details). Therefore, in this case (corresponding to a simple real characteristic value λ_0) we have

$$\mu'(\lambda_0) = -\frac{E_\lambda(\lambda_0;0)}{E_\mu(\lambda_0;0)}, \tag{4.9}$$

and then comparing with the graphical formula (1.9) for the Krein signature of such a characteristic value we learn that

$$\kappa(\lambda_0) = \operatorname{sign}(\mu'(\lambda_0)) = -\operatorname{sign}\left[\frac{E_\lambda(\lambda_0;0)}{E_\mu(\lambda_0;0)}\right]. \tag{4.10}$$

Recalling that $E(\lambda;0)$ is necessarily an Evans function $D(\lambda)$ for $\mathcal{L}$, this formula can be re-written in the equivalent form

$$\kappa(\lambda_0) = -\operatorname{sign}\left[\frac{D'(\lambda_0)}{E_\mu(\lambda_0;0)}\right]. \tag{4.11}$$

This formula[12] shows how the Krein signature of a simple real characteristic value λ_0 indeed relates to the derivative of an Evans function; the reason that the relationship seems ambiguous as shown in Examples 9–10 is that to obtain agreement one must take into account an additional factor, namely the sign of $E_\mu(\lambda_0;0)$ which can be different for different characteristic values λ_0.

[12]The specific Evans function $D(\lambda)$ in the numerator of course depends on the Evans-Krein function $E(\lambda;\mu)$ from which it arises upon setting $\mu = 0$. This has some immediate implications; for example if one considers a standard eigenvalue problem $Lu = \lambda u$ for a selfadjoint operator L, upon introducing the selfadjoint pencil $\mathcal{L}(\lambda) := L - \lambda I$ and the corresponding μ-perturbation $\mathcal{M}(\lambda;\mu) := L - \lambda I - \mu I = L - (\lambda + \mu)I$ one sees that $E_\lambda(\lambda_0;0) = E_\mu(\lambda_0;0)$ for all real characteristic values (eigenvalues of L) λ_0. Hence the Krein signature of each simple characteristic value is necessarily positive, even though by Rolle's Theorem $D'(\lambda) = E_\lambda(\lambda;0)$ will have sign changes.

An Evans-Krein function $E(\lambda;\mu)$ can also be used to provide information about the positive and negative Krein indices of the eigenvalue branch(es) passing through $\mu = 0$ at an isolated non-simple real characteristic value λ_0 of a selfadjoint operator pencil. According to the theory presented in §3 (see Definition 3.3), the (graphical) Krein indices $\kappa_{\rm g}^{\pm}(\mu,\lambda_0)$ associated with the eigenvalue branch $\mu = \mu(\lambda)$ are known in the case of odd algebraic multiplicity (order of vanishing of $\mu(\lambda)$ at λ_0) m once the sign of $\mu^{(m)}(\lambda_0) \neq 0$ is known (if m is even there is nothing to calculate). But $\mu^{(m)}(\lambda_0)$ can also be expressed in terms of nonzero partial derivatives of an Evans-Krein function $E(\lambda;\mu)$ at the point $(\lambda,\mu) = (\lambda_0,0)$. The main idea is to calculate sufficiently many derivatives of the identity $E(\lambda;\mu(\lambda)) = 0$ at the point $\lambda = \lambda_0$. By repeated application of the chain rule, one finds that

$$\frac{d^r}{d\lambda^r}E(\lambda;\mu(\lambda)) = \sum_{n=0}^{r}\binom{r}{n}\frac{\partial^r E}{\partial\mu^n\partial\lambda^{r-n}}(\lambda;\mu(\lambda))\mu'(\lambda)^n$$
$$+\sum_{p=1}^{r-1}\sum_{n+s=p}\sum_{\substack{d_1\le\cdots\le d_n\\ d_1+\cdots+d_n=p}}\beta_{r,n,s}(d_1,\ldots,d_n)\frac{\partial^{r-s}E}{\partial\mu^n\partial\lambda^{r-p}}(\lambda;\mu(\lambda))\prod_{j=1}^{n}\mu^{(d_j)}(\lambda), \quad (4.12)$$

where n, s, and d_j for $j = 1,\ldots,n$ are positive integers and $\beta_{r,n,s}(d_1,\ldots,d_n)$ are certain complicated coefficients. Many of the partial derivatives of E can be evaluated at $\lambda = \lambda_0$ and $\mu(\lambda_0) = 0$ by using the normal form (4.6). While one can develop a general theory, to keep things simple we just present two representative examples.

◁ *Example 11. Real characteristic values of geometric multiplicity $k = 1$.* For a real characteristic value λ_0 of $\mathcal{L}$ having geometric multiplicity $k = 1$ but arbitrary finite algebraic multiplicity $m \ge 1$, there is a unique eigenvalue branch $\mu = \mu_1(\lambda)$ (analytic under suitable assumptions) of the problem (3.1) passing through the point $(\lambda,\mu) = (\lambda_0,0)$, and we have $\mu_1(\lambda_0) = \mu_1'(\lambda_0) = \cdots = \mu_1^{(m-1)}(\lambda_0) = 0$ but $\mu_1^{(m)}(\lambda_0) \neq 0$. If m is even, by definition we have $\kappa_{\rm g}^{\pm}(\mu_1,\lambda_0) = m/2$, while for m odd we need to calculate $\mu_1^{(m)}(\lambda_0)$ to determine the Krein indices. But this is easily obtained from an Evans-Krein function $E(\lambda;\mu)$ by calculating m derivatives of the identity $E(\lambda;\mu_1(\lambda)) = 0$ using (4.12) and evaluating at $\lambda = \lambda_0$. Only two terms survive for $\lambda = \lambda_0$ due to the fact that $\mu_1(\lambda)$ vanishes there to order m. Indeed,

$$\left.\frac{d^m}{d\lambda^m}E(\lambda;\mu_1(\lambda))\right|_{\lambda=\lambda_0} = \frac{\partial^m E}{\partial\lambda^m}(\lambda_0;0) + \frac{\partial E}{\partial\mu}(\lambda_0;0)\mu_1^{(m)}(\lambda_0) = 0 \quad (4.13)$$

and since $E_\mu(\lambda_0;0) \neq 0$ as follows from the fact that $k = 1$, we obtain

$$\mu_1^{(m)}(\lambda_0) = -\frac{1}{E_\mu(\lambda_0;0)}\frac{\partial^m E}{\partial\lambda^m}(\lambda_0;0). \quad (4.14)$$

Note that $\partial^m E/\partial\lambda^m(\lambda_0;0) \neq 0$ because this is the m-th derivative of an Evans function $D(\lambda)$ at a characteristic value of algebraic multiplicity m. ▷

◁ *Example 12. Semisimple real characteristic values.* If λ_0 is a semisimple real characteristic value of a selfadjoint operator pencil $\mathcal{L}$ of geometric and algebraic multiplicity $k \ge 1$, then there are k (analytic under suitable assumptions) eigenvalue branches $\mu = \mu_j(\lambda)$, $j = 1,\ldots,k$, of the problem (3.1) each of which has a simple root at $\lambda = \lambda_0$. To determine the Krein indices it suffices to express $\mu_j'(\lambda_0)$ in terms

of an Evans-Krein function $E(\lambda;\mu)$ for $j=1,\dots,k$. Now, from the normal form (4.6) one can easily see that in this case $E(\lambda_0;\mu) = C(\lambda_0;\mu)\mu^k$ with $C(\lambda_0;0) \neq 0$ so in particular $E_\mu(\lambda_0;0) = 0$ unless $k=1$ (and then λ_0 is simple), implying that in general the method of differentiation once along each branch individually and evaluating at $\lambda=\lambda_0$ as in the calculation leading to (4.9) will fail.

However, using the normal form (4.6) and the first-order vanishing of all k of the branches crossing at $(\lambda_0,0)$ one can easily show that

$$\frac{\partial^{n+j}E}{\partial\mu^n\partial\lambda^j}(\lambda_0;0) = 0, \quad n+j<k. \tag{4.15}$$

If we choose a branch $\mu=\mu_j(\lambda)$ and differentiate the identity $E(\lambda;\mu_j(\lambda))=0$ exactly k times with respect to λ at $\lambda=\lambda_0$, then with $r=k$ in (4.12) the identities (4.15) guarantee that all of the terms on the second line vanish identically. Therefore the k values $z=\mu_j'(\lambda_0)$ are determined as the roots of the k-th order polynomial equation

$$\sum_{n=0}^{k}\binom{k}{n}\frac{\partial^k E}{\partial\mu^n\partial\lambda^{k-n}}(\lambda_0;0)z^n = 0. \tag{4.16}$$

This equation is genuinely k-th order because the normal form (4.6) and our assumptions on $\mu_j(\lambda)$ imply that

$$\frac{\partial^k E}{\partial\mu^n\partial\lambda^{k-n}}(\lambda_0;0) = n!C(\lambda_0;0)\frac{d^{k-n}\gamma_n}{d\lambda^{k-n}}(\lambda_0). \tag{4.17}$$

Here, $\gamma_p(\lambda)$ is defined as the coefficient of μ^p in the product

$$\prod_{j=1}^{k}(\mu-\mu_j(\lambda)) = \sum_{p=0}^{k}\gamma_p(\lambda)\mu^p. \tag{4.18}$$

Since $\gamma_k(\lambda)\equiv 1$, the coefficient of z^k in (4.16) is $k!C(\lambda_0;0)\neq 0$. Likewise, since $\gamma_0(\lambda) = (-1)^k\mu_1(\lambda)\cdots\mu_k(\lambda)$, the constant term in the polynomial on the left-hand side of (4.16) is $(-1)^kC(\lambda_0;0)\mu_1'(\lambda_0)\cdots\mu_k'(\lambda_0)\neq 0$, which implies that all of the roots of (4.16) are nonzero. That all k of the roots are real numbers is less obvious but necessarily true. ▷

In summary, as was observed in [36, 38], an Evans function $D(\lambda)$ alone cannot explain the origin of Krein indices and signatures of characteristic values. The geometrical reason is simple: $D(\lambda)$ is proportional to the *product of all eigenvalue branches* $\mu(\lambda)$ of the problem (3.1), and no amount of differentiation with respect to λ will separate the individual branches. However, the branches can indeed be separated with the help of derivatives of an Evans-Krein function $E(\lambda;\mu)$ with respect to μ at $(\lambda,\mu)=(\lambda_0,0)$ that encode the local behavior of each branch close to the characteristic value.

4.5. Further generalization of Evans-Krein functions. For some purposes (see Example 13 below) it may be important to determine whether two distinct real characteristic values of a selfadjoint operator pencil $\mathcal{L}$ belong to (are obtained as roots of) the "same" eigenvalue branch $\mu=\mu(\lambda)$ of the associated eigenvalue problem (3.1). Determining whether two real characteristic values are connected in this way requires analyzing the eigenvalue branches of the operator $\mathcal{L}(\lambda)$ away from $\mu=0$. When we

described a typical construction of an Evans-Krein function in problems of stability analysis for nonlinear waves, we pointed out that when μ is sufficiently small the hyperbolicity of the asymptotic systems (4.4) is guaranteed if it is present for $\mu = 0$. If we wish to have a construction that works for larger values of μ, we can generalize the definition of an Evans-Krein function in the following way.

Instead of (3.1) one can consider the μ-deformed characteristic value problem

$$\mathcal{L}(\lambda)u = \mu S u, \tag{4.19}$$

and define a generalized Evans-Krein function as an Evans function for the pencil $\mathcal{M}_S(\lambda;\mu) := \mathcal{L}(\lambda) - \mu S$, where S is a suitable positive definite selfadjoint operator on X. The correspondence between traditional and graphical Krein indices/signatures given in Theorem 3.4 can be shown to hold when the eigenvalue branches are defined by the modified problem (4.19) for a wide class of positive definite selfadjoint operators $S : X \to X$. In the case of spectral analysis of nonlinear wave equations linearized about a localized traveling wave solution (for which the Evans function construction described in §4.1 applies), it is enough to define $Su(x) = f(x)u(x)$ where $f(x)$ is a positive Schwartz-class function. The rapid decay of f as $|x| \to \infty$ guarantees that for all μ however large the asymptotic systems (4.4) are exactly the same as they are for $\mu = 0$ and hence hyperbolicity is preserved. Such a generalized Evans-Krein function should be straightforward to implement in existing computer codes for evaluating Evans functions constructed as described in §4.1. Finally, we can observe that the freedom of choice of the operator S in (4.19) suggests a certain structural stability of the problem of determining Krein indices and signatures of real characteristic values of the problem (2.23). A specific example of a kind of problem in which one needs to consider values of μ that are not small is the following.

$\lhd$ *Example 13. Stability of vortices in axisymmetric Bose-Einstein condensates.* The wave function $\Psi = \Psi(x,t)$ of a Bose-Einstein condensate confined in a harmonic trap evolves according to the Gross-Pitaevskii equation

$$i\hbar\psi_t = \left(-\frac{1}{2}\Delta + \frac{|x|^2}{2} + |\psi|^2\right)\psi, \tag{4.20}$$

where Δ is the Laplacian. In the two-dimensional case ($x \in \mathbb{R}^2$, $|x|^2 := x_1^2 + x_2^2$) the Gross-Pitaevskii equation (4.20) has solitary wave (vortex) solutions of the form

$$\psi(t,r,\theta) = e^{-i(m+1+\mu)t}e^{im\theta}w_{\mu,m}(r), \tag{4.21}$$

where (r,θ) are polar coordinates, m is the *degree* of the vortex, $m+1+\mu$ is the vortex rotation frequency and $w_{\mu,m}(r)$ is the radial vortex profile.

The stability of vortices in Bose-Einstein condensates has been well-studied. For example, in [44] an Evans function was used to study the spectral stability of vortices of degree $m = 1, 2$ which exist for $\mu \geq 0$. Reformulating the spectral stability problem studied in [44] in the language of selfadjoint operator pencils, spectral instability corresponds to the presence of non-real characteristic values (mode frequencies) λ of the spectral problem $\mathcal{L}_{\mathrm{BEC}}(\lambda)y = 0$ for a function $y = (y_+(r), y_-(r))^{\mathsf{T}} \in L^2(\mathbb{R}_+, \mathbb{C}^2; r\,dr)$, where $\mathcal{L}_{\mathrm{BEC}}(\lambda) = \lambda L_1 + L_0$ is the linear pencil with coefficients

$$\begin{aligned} L_1 &:= -\sigma_3, \\ L_0 &:= \left(-\frac{1}{2r}\frac{d}{dr}r\frac{d}{dr} + \frac{1}{2}r^2 - m - 1 - \mu + \frac{j^2+m^2}{2r^2} + 2|w_{\mu,m}(r)|^2\right)\mathbb{I} \\ &\quad + \frac{jm}{r^2}\sigma_3 + |w_{\mu,m}(r)|^2\sigma_1. \end{aligned} \tag{4.22}$$

Here, j is an integer satisfying $0 < |j| < 2m$ that indexes an angular Fourier mode. Note that in the indicated weighted L^2 space, this operator pencil $\mathcal{L}_{\rm BEC}$ is selfadjoint for bounded y that decay rapidly as $r \to \infty$.

In [44] the radial profile of the vortex and then the corresponding characteristic values of $\mathcal{L}_{\rm BEC}$ were calculated numerically for a moderately large range of values of $\mu \geq 0$. Note that $\mu \geq 0$ enters into the pencil $\mathcal{L}_{\rm BEC}$ both through the linear term $-\mu\mathbb{I}$ and also through the radial profile $|w_{\mu,m}(r)|^2$ of the vortex.

For $m = 1$ all the characteristic values of $\mathcal{L}_{\rm BEC}$ are real and thus the vortex is spectrally stable, while for $m = 2$ variation of the parameter μ leads to Hamiltonian-Hopf bifurcations creating "bubbles" of instability [55] in which two colliding characteristic values leave and subsequently return to the real axis. As it is typical in Hamiltonian systems, the underlying Hamiltonian has only a few negative eigenvalues limiting the number of possible real characteristic values of negative Krein signature. In fact, for unstable vortices ($m = 2$, $j = 2$) there is only one real characteristic value of negative Krein signature that repeatedly collides with the remaining real characteristic values of positive signature as the parameter μ increases. In [44] a rigorous theory of continuity of Krein signatures for finite systems of characteristic values was developed, and its results were used to identify the unique real characteristic value having negative Krein signature via a continuation (homotopy) method in μ starting from the explicitly solvable case of $\mu = 0$.

The apparent weakness of this method is that a stability check for large values of μ requires a significant computational overhead — the calculation of the characteristic values of the pencil $\mathcal{L}_{\rm BEC}$ for a discrete sampling of a large interval of μ that is sufficiently dense to ensure continuity of characteristic values (and their Krein signatures). Naturally, once a characteristic value λ has been calculated for any value of μ it is possible to determine its Krein signature by an evaluation of the particular quadratic form as in (1.8) at the corresponding characteristic vector. Unfortunately, the Evans function used in [44] was constructed from exterior products, and consequently it is not obvious how to numerically recover the characteristic vectors from the Evans function; this meant that the characteristic vectors and subsequently the Krein signature of each characteristic value had to be calculated separately. By contrast, it would be easy to capture the Krein signature of any real characteristic value directly from an Evans-Krein function with a minimal extra computational and coding cost for any value of μ, thus making calculations for large μ directly accessible. $\triangleright$

5. Index Theorems for Linearized Hamiltonian Systems. The graphical interpretation of Krein indices and signatures afforded by Theorem 3.4 can be applied to develop very simple proofs of some well-known *index theorems* (eigenvalue counts, inertia laws) for linearized Hamiltonian systems that appear frequently in stability analysis [24, 26, 37, 65] where the central question is the presence of *unstable spectrum* of (1.1) in the right-half complex plane. Our goal is to show that the counts of unstable spectrum provided by these index theorems can be interpreted in terms of topological indices of planar curves in a quadrant.

Recall Example 1, in which it was shown that unstable spectrum can only be present if L is indefinite. In applications to stability of nonlinear waves L typically represents the linearization of a Hamiltonian whose kinetic energy part is a positive-definite unbounded differential operator. In such cases L will itself be unbounded in the positive direction but may have a finite number of negative eigenvalues that might cause instability of the system. In this context, the purpose of an index theorem is to bound the number of possible points in the unstable (non-imaginary) part of the

spectrum $\sigma(JL)$ of (1.1) in terms of the total dimension of negative eigenspaces of L and information about the stable (purely imaginary) part of the spectrum. It turns out that this information amounts to the number of real characteristic values λ of the corresponding linear pencil $\mathcal{L}$ given by (1.3) having certain types of Krein signatures. (Recall that the real characteristic values of $\mathcal{L}$ are in one-to-one correspondence with the purely imaginary points ν of $\sigma(JL)$ simply by a rotation of the complex plane: $\lambda = i\nu$.) Such theorems therefore generalize the simple statement that positivity of L prevents instability. An important application of index theorems is the justification of a numerical search for unstable spectrum; once the number of complex characteristic values of $\mathcal{L}$ obtained numerically (say via Argument Principle calculations involving an appropriate Evans function) equals the maximum number admitted by the index theorem, the numerical search can be considered complete [44].

For the sake of clarity, we will work exclusively in the finite-dimensional setting. All of our applications below will be based on Theorem 5.2, and while some aspects of its proof are not difficult to generalize to the infinite-dimensional setting (for instance, by counting codimensions rather than dimensions), the possible presence of an infinite number of discrete eigenvalue branches complicates other aspects of the proof. This subject will be addressed in a separate paper.

To explain the index theorems graphically, we shift the focus from the local information contained in the way the eigenvalue branches $\mu(\lambda)$ of (1.5) cross $\mu = 0$ to the global topological information stored in the branches for μ not small.

5.1. Graphical analysis of selfadjoint polynomial matrix pencils. We begin by formulating some simple consequences of graphical analysis of eigenvalue curves $\mu = \mu(\lambda)$ corresponding to selfadjoint polynomial matrix pencils of arbitrary (odd) degree. In §5.2 below we will apply these results in the special case of linear pencils, but it is easy to envision future applications where greater generality could be useful. We use the following notation below.

DEFINITION 5.1. *If L is a Hermitian $N \times N$ matrix, the number of strictly positive (respectively negative) eigenvalues of L counted with multiplicity i s denoted $N_+(L)$ (respectively $N_-(L)$).*

If $\mathcal{L}$ is a selfadjoint $N \times N$ matrix pencil, the number of strictly positive (respectively negative) real characteristic values of $\mathcal{L}$ counted with geometric multiplicity is denoted $N_+(\mathcal{L})$ (respectively $N_-(\mathcal{L})$). The number of eigenvalue curves $\mu = \mu(\lambda)$ of $\mathcal{L}$ passing through $(\lambda, \mu) = (0, 0)$ with $\mu(\lambda) < 0$ for sufficiently small positive (respectively negative) λ, will be denoted $Z_+^{\downarrow}(\mathcal{L})$ (respectively $Z_-^{\downarrow}(\mathcal{L})$), and analogously the number of eigenvalue curves $\mu = \mu(\lambda)$ for which $\mu(\lambda) > 0$ for sufficiently small positive (respectively negative) λ will be denoted $Z_+^{\uparrow}(\mathcal{L})$ (respectively $Z_-^{\uparrow}(\mathcal{L})$).

Note that whenever we refer to a number of eigenvalue curves as in Definition 5.1, we intend them to be counted weighted by their geometric multiplicity. It is important to stress that there may be curves counted in *both* $Z_+^{\downarrow}(\mathcal{L})$ and $Z_-^{\downarrow}(\mathcal{L})$, namely those vanishing at $\lambda = 0$ to even order m with $\mu^{(m)}(0) < 0$. If the order of vanishing m is even but $\mu^{(m)}(0) > 0$, then the corresponding curve will be counted in *neither* $Z_+^{\downarrow}(\mathcal{L})$ nor $Z_-^{\downarrow}(\mathcal{L})$. Curves vanishing at $\lambda = 0$ to odd order will be counted in exactly one of $Z_+^{\downarrow}(\mathcal{L})$ or $Z_-^{\downarrow}(\mathcal{L})$.

Of course $N_+(L) + N_-(L) + \dim(\mathrm{Ker}(L)) = N$, and also $Z_-^{\downarrow}(\mathcal{L}) + Z_-^{\uparrow}(\mathcal{L}) = Z_+^{\downarrow}(\mathcal{L}) + Z_+^{\uparrow}(\mathcal{L}) = \dim(\mathrm{Ker}(\mathcal{L}(0)))$. In general neither $N_+(\mathcal{L})$ nor $N_-(\mathcal{L})$ is necessarily finite, although the situation is better for polynomial pencils as we now recall. A polynomial matrix pencil $\mathcal{L} = \mathcal{L}(\lambda) := L_0 + \lambda L_1 + \cdots + \lambda^p L_p$ acting in $X = \mathbb{C}^N$

equipped (as will be understood throughout this section) with the Euclidean inner product is selfadjoint if and only if the coefficients $L_0, \ldots, L_p$ are all represented in a fixed basis of X as Hermitian $N \times N$ matrices. According to Definition 2.1, the total algebraic multiplicity of all characteristic values making up the spectrum $\sigma(\mathcal{L})$ of such a pencil is equal to pN under the additional assumption that the leading coefficient L_p is an invertible matrix. We have the following result.

THEOREM 5.2. *Let $\mathcal{L} = \mathcal{L}(\lambda) := L_0 + \lambda L_1 + \cdots + \lambda^p L_p$ be a selfadjoint matrix pencil of odd degree $p = 2\ell + 1$ acting on $X = \mathbb{C}^N$, and suppose that L_p is invertible. Then we have the fundamental graphical conservation law*

$$N - 2N_-(L_0) - Z_+^{\downarrow}(\mathcal{L}) - Z_-^{\downarrow}(\mathcal{L}) + \sum_{\substack{\lambda \in \sigma(\mathcal{L}) \\ \lambda > 0}} \kappa(\lambda) - \sum_{\substack{\lambda \in \sigma(\mathcal{L}) \\ \lambda < 0}} \kappa(\lambda) = 0. \tag{5.1}$$

Also, the following inequalities hold true:

$$N_{\pm}(\mathcal{L}) \geq \left| N_-(L_0) + Z_{\pm}^{\downarrow}(\mathcal{L}) - N_{\mp}(L_p) \right|. \tag{5.2}$$

Proof. According to Theorem 2.2, the N eigenvalue curves $\mu = \mu(\lambda)$ associated with the selfadjoint pencil $\mathcal{L}$ may be taken as holomorphic functions of $\lambda \in \mathbb{R}$. The assumption that L_p is invertible implies that none of the eigenvalue curves can be a constant function, and hence each curve has a well-defined finite order of vanishing at each corresponding real characteristic value of $\mathcal{L}$. Let $Q_{\pm}$ denote the open quadrants of the (λ, μ)-plane corresponding to $\mu < 0$ and $\pm\lambda > 0$. By analyticity and hence continuity of the eigenvalue branches $\mu = \mu(\lambda)$, the number of branches entering $Q_{\pm}$ necessarily equals the number of branches leaving $Q_{\pm}$ (with increasing λ).

Consider first the quadrant Q_-. Curves $\mu = \mu(\lambda)$ can enter Q_- with increasing λ in only two ways: from $\lambda = -\infty$ (corresponding to curves with $\mu(\lambda) < 0$ asymptotically as $\lambda \to -\infty$) and through the negative λ-axis. Since the odd-degree polynomial pencil $\mathcal{L}$ is dominated by its invertible leading term $\lambda^p L_p$ as $|\lambda| \to \infty$ (as is easy to see by applying Rouché's Theorem to the characteristic polynomial of the matrix $\lambda^{-p}\mathcal{L}(\lambda)$) , the number of curves entering Q_- from $\lambda = -\infty$ is precisely $N_+(L_p)$. Similarly, curves $\mu = \mu(\lambda)$ can exit Q_- with increasing λ in exactly three ways: through the negative μ-axis, through the origin $(\lambda, \mu) = (0, 0)$, and through the negative λ-axis. Since $\mathcal{L}(0) = L_0$, the number of curves exiting Q_- through $\mu < 0$ is precisely $N_-(L_0)$, and $Z_-^{\downarrow}(\mathcal{L})$ is the count of curves exiting Q_- through the origin. Therefore "conservation of curves" for the quadrant Q_- reads

$$\begin{aligned} N_+(L_p) - N_-(L_0) - Z_-^{\downarrow}(\mathcal{L}) \\ = \text{Net number of curves exiting } Q_- \text{ through } \lambda < 0, \mu = 0. \end{aligned} \tag{5.3}$$

If λ_0 is a real characteristic value of $\mathcal{L}$, then according to Definition 3.3 and Theorem 3.4, the Krein signature $\kappa(\lambda_0)$ exactly counts the net number of eigenvalue branches $\mu = \mu(\lambda)$ vanishing at λ_0 exiting the half-plane $\mu < 0$ with increasing λ. Therefore conservation of curves for Q_- can be rewritten equivalently as

$$N_+(L_p) - N_-(L_0) - Z_-^{\downarrow}(\mathcal{L}) = \sum_{\lambda \in \sigma(\mathcal{L}), \lambda < 0} \kappa(\lambda). \tag{5.4}$$

A conservation law for curves in Q_+ is obtained similarly:

$$\begin{aligned} N_-(L_p) - N_-(L_0) - Z_+^{\downarrow}(\mathcal{L}) \\ &= \text{Net number of curves entering } Q_+ \text{ through } \lambda > 0,\ \mu = 0 \\ &= -\sum_{\lambda\in\sigma(\mathcal{L}),\lambda>0} \kappa(\lambda). \end{aligned} \tag{5.5}$$

Adding together (5.4) and (5.5), and taking into account the fact that $N_-(L_p) + N_+(L_p) = N$ because L_p is Hermitian $N\times N$ and invertible, we arrive at the graphical conservation law (5.1).

The inequalities (5.2) follow immediately from the equations (5.3) and (5.5) by noting that the absolute value of the net number of curves crossing the positive or negative λ-semiaxis is a lower bound for the number of curves crossing the same axis, and the latter is exactly a count of real characteristic values according to their geometric multiplicity. □

The key ideas of the proof are illustrated in Figure 5.1. As the proof is essentially

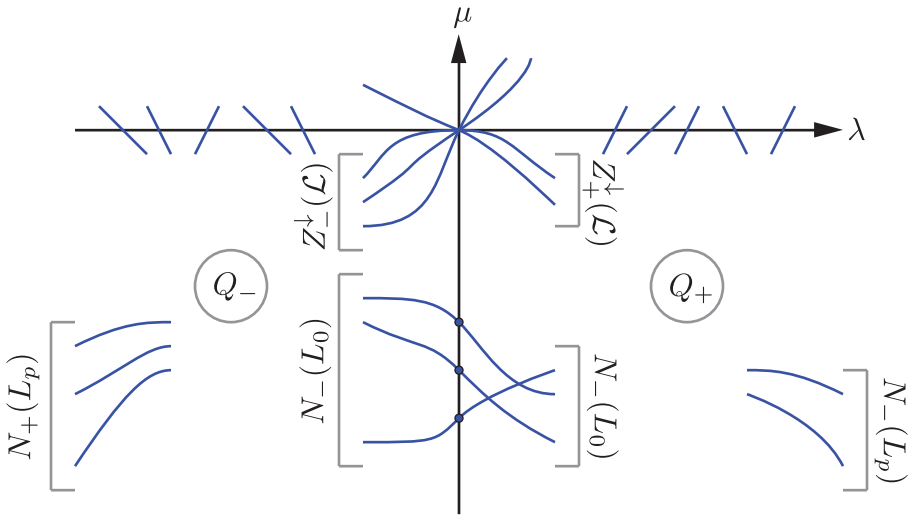

FIG. 5.1. *The proof of Theorem 5.2 is essentially graphical in nature. In each quadrant $Q_\pm$ the number of curves entering and exiting through infinity, the negative μ-axis, and through the origin must be balanced by the net numbers exiting through the positive and negative λ-semiaxes as calculated by summing Krein signatures. The negative eigenvalues of L_0 are plotted with blue dots.*

topological in nature, the analyticity of the branches implied by Theorem 2.2 could be replaced by mere continuity.

5.2. Application to linearized Hamiltonian systems. We now return to (1.1), the spectral problem $JLu = \nu u$ for a linearized Hamiltonian system, where L and J are (Hermitian and invertible skew-Hermitian, respectively) matrices of size $N \times N$, and we recall that in the finite-dimensional setting the underlying space $X = \mathbb{C}^N$ necessarily has even dimension $N = 2n$ due to invertibility of the skew-Hermitian matrix J. Let the total algebraic multiplicity of all points $\nu \in \sigma(JL)$ with $\mathrm{Re}\{\nu\} \neq 0$ be denoted $2n_{\mathrm{u}}$ (necessarily an even number due to the basic Hamiltonian

symmetry $\nu \to -\overline{\nu}$). This is the dimension of the *unstable invariant subspace* of JL. The total algebraic multiplicity of all purely imaginary points of $\sigma(JL)$ is therefore $2n - 2n_{\rm u} = 2n_{\rm s}$, so $2n_{\rm s}$ is the dimension of the *(spectrally) stable invariant subspace* of JL. The first result is the following.

THEOREM 5.3 (Index theorem for linearized Hamiltonians). *Let L and J be $2n \times 2n$ matrices, and let L be Hermitian and J be invertible and skew-Hermitian. Let $\mathcal{L} = \mathcal{L}(\lambda) := L - \lambda K$, $K := (iJ)^{-1}$, be the associated linear matrix pencil. Then*

$$n_{\rm u} = N_-(L) - \zeta - \sum_{\lambda \in \sigma(\mathcal{L}), \lambda > 0} \kappa^+(\lambda) - \sum_{\lambda \in \sigma(\mathcal{L}), \lambda < 0} \kappa^-(\lambda). \tag{5.6}$$

The Krein indices $\kappa^{\pm}(\lambda) \geq 0$ are associated with real characteristic values λ of the selfadjoint pencil $\mathcal{L}$, which in turn correspond via rotation of the complex plane to purely imaginary points in $\sigma(JL)$. Here the quantity ζ is given by

$$\zeta = \frac{1}{2}\dim(\mathrm{gKer}(JL)) - \frac{1}{2}\left(Z_+^{\downarrow}(\mathcal{L}) + Z_-^{\downarrow}(\mathcal{L})\right), \tag{5.7}$$

or equivalently by

$$\zeta := \frac{1}{2}\left(\dim(\mathrm{gKer}(JL)) - 2\dim(\mathrm{Ker}(L))\right) + \frac{1}{2}\left(Z_+^{\uparrow}(\mathcal{L}) + Z_-^{\uparrow}(\mathcal{L})\right). \tag{5.8}$$

Proof. We apply Theorem 5.2 to the selfadjoint linear pencil $\mathcal{L}$ associated with L and J. Here $p = 1$, and $L_0 = L$ while $L_1 = -K$. Note that K is invertible. With $N = 2n = 2n_{\rm u} + 2n_{\rm s}$, the graphical conservation law (5.1) reads (upon rearrangement and division by 2)

$$n_{\rm u} = N_-(L) + \frac{1}{2}(Z_+^{\downarrow}(\mathcal{L}) + Z_-^{\downarrow}(\mathcal{L})) - n_{\rm s} - \frac{1}{2}\sum_{\substack{\lambda \in \sigma(\mathcal{L}) \\ \lambda > 0}} \kappa(\lambda) + \frac{1}{2}\sum_{\substack{\lambda \in \sigma(\mathcal{L}) \\ \lambda < 0}} \kappa(\lambda). \tag{5.9}$$

Of course by definition the Krein signatures $\kappa(\lambda)$ appearing on the right-hand side can be written as differences of two positive Krein indices:

$$\kappa(\lambda) = \kappa^+(\lambda) - \kappa^-(\lambda). \tag{5.10}$$

But $n_{\rm s}$ can also be written in terms of the Krein indices of the real characteristic values of $\mathcal{L}$. Indeed, the total algebraic multiplicity of a purely imaginary point $\nu \in \sigma(JL)$ is equal to that of the corresponding real characteristic value $\lambda = i\nu$ of $\mathcal{L}$, and the latter can be expressed as the sum of the Krein indices $\kappa^+(\lambda) + \kappa^-(\lambda)$. Therefore, summing over all real characteristic values of $\mathcal{L}$ we obtain:

$$2n_{\rm s} = \sum_{\lambda \in \sigma(\mathcal{L}), \lambda \in \mathbb{R}} \left(\kappa^+(\lambda) + \kappa^-(\lambda)\right). \tag{5.11}$$

Substituting this identity and (5.10) into (5.9), and noting that as a count of algebraic multiplicity we have $\kappa^+(0) + \kappa^-(0) = \dim(\mathrm{gKer}(JL))$ if $0 \in \sigma(JL)$, we obtain (5.6) with ζ given in the form (5.7). Since $Z_+^{\uparrow}(\mathcal{L}) + Z_+^{\downarrow}(\mathcal{L}) = Z_-^{\uparrow}(\mathcal{L}) + Z_-^{\downarrow}(\mathcal{L}) = \dim(\mathrm{Ker}(\mathcal{L}(0))) = \dim(\mathrm{Ker}(L))$ the alternative formula (5.8) follows immediately. □

Recall that if J and L are real matrices, then in addition to the Hamiltonian symmetry $\sigma(JL) = -\overline{\sigma(JL)}$ we have the real symmetry $\sigma(JL) = \overline{\sigma(JL)}$, in which case

system is said to have *full Hamiltonian symmetry*. In many applications great utility is gleaned from this extra symmetry, but one should be aware that the reality of the matrices J and L (or operators in the infinite-dimensional setting) is essentially tied to a choice of basis or representation. In other words, within the usual Hilbert space axioms one finds no inherent notion of what it means for a vector or operator to be "real". Therefore, a problem can possess full Hamiltonian symmetry even though it may not be obvious that the operators are "real" due to poor choice of coordinates. A useful general notion of reality is based instead on the existence of an appropriate antiholomorphic involution on X and can be formulated as follows (compare with [29, Assumption 2.15]).

DEFINITION 5.4. *Let $I : X \to X$ be an involution on a Hilbert space X that is unitary: $\|I(u)\| = \|u\|$, $\forall u \in X$, and conjugate linear: $I(\alpha u + \beta v) = \overline{\alpha} I(u) + \overline{\beta} I(v)$, $\forall u, v \in X$ and $\forall \alpha, \beta \in \mathbb{C}$. A linear operator $A : X \to X$ is said to be real (with respect to I) if it commutes with I: $AI(u) = I(Au)$ for all $u \in X$. Similarly, a vector $u \in X$ is real (with respect to I) if $u = I(u)$.*

The involution I is simply an abstraction of complex conjugation as might be applied to complex coordinates of vectors in X. In the case of real matrices J and L we have the following important corollary of Theorem 5.3.

COROLLARY 5.5 (Index theorem for linearized Hamiltonians with full Hamiltonian symmetry). *Let L and J be $2n \times 2n$ matrices, and let L be Hermitian and J be invertible and skew-Hermitian. Suppose also that L and J are real with respect to a unitary antiholomorphic involution $I : \mathbb{C}^{2n} \to \mathbb{C}^{2n}$. Then with ζ being given by either* (5.7) *or* (5.8)*,*

$$n_{\mathrm{u}} = N_-(L) - \zeta - 2 \sum_{\lambda \in \sigma(\mathcal{L}), \lambda > 0} \kappa^+(\lambda) = N_-(L) - \zeta - 2 \sum_{\lambda \in \sigma(\mathcal{L}), \lambda < 0} \kappa^-(\lambda), \tag{5.12}$$

and moreover, ζ can be simplified as follows:

$$\begin{aligned} \zeta &= \frac{1}{2} \dim(\mathrm{gKer}(JL)) - Z^-(\mathcal{L}) \\ &= \frac{1}{2} \left(\dim(\mathrm{gKer}(JL)) - 2 \dim(\mathrm{Ker}(L)) \right) + Z^+(\mathcal{L}), \end{aligned} \tag{5.13}$$

where $Z^+(\mathcal{L})$ (respectively $Z^-(\mathcal{L})$) is the number of curves $\mu = \mu(\lambda)$ vanishing at $\lambda = 0$ whose first nonzero derivative $\mu^{(m)}(\lambda)$ is positive (respectively negative). If in addition one has $\dim(\mathrm{gKer}(JL)) = 2 \dim(\mathrm{Ker}(L))$, then $\zeta = Z^+(\mathcal{L})$.

Proof. The fact that JL commutes with the involution I forces the whole system of curves $\mu = \mu(\lambda)$ to be symmetric with respect to reflection through the vertical μ-axis. This implies that if $\lambda_0 > 0$ is a real characteristic value (with root space U) of the associated linear pencil $\mathcal{L}$, then so is $-\lambda_0$ (with root space $I(U)$), and from the definition of graphical Krein indices (see Definition 3.3 and Theorem 3.4) we see that $\kappa^{\pm}(\lambda_0) = \kappa^{\mp}(-\lambda_0)$ due to the symmetry of the curves. Therefore, both sums on the right-hand side of (5.6) are equal, and (5.12) follows immediately.

The simplified formulae (5.13) for ζ follow from the left-right symmetry of the union of curves $\mu = \mu(\lambda)$ passing through the origin $(\lambda, \mu) = (0, 0)$. Indeed, the symmetry implies that the number of curves vanishing to odd order m is even and divides equally into the number of those curves for which $\mu^{(m)}(0) > 0$ and that of their reflections through the μ-axis. This implies that $Z^{\downarrow}_+(\mathcal{L}) + Z^{\downarrow}_-(\mathcal{L}) = 2Z^-$, and then since $\dim(\mathrm{Ker}(L)) = \dim(\mathrm{Ker}(JL))$ is the total number of analytic branches

through the origin, the proof of (5.13) is complete. The fact that the condition $\dim(\mathrm{gKer}(JL)) = 2\dim(\mathrm{Ker}(L))$ implies that $\zeta = Z^+(\mathcal{L})$ is obvious. □

The next application concerns spectral problems for linearized Hamiltonian systems for which L and J have so-called *canonical form*, that is, the $2n \times 2n$ matrices can be written in terms of $n \times n$ blocks as follows:

$$J = \begin{pmatrix} 0 & \mathbb{I} \\ -\mathbb{I} & 0 \end{pmatrix}, \quad L = \begin{pmatrix} L_+ & 0 \\ 0 & L_- \end{pmatrix} \tag{5.14}$$

where $\mathbb{I}$ is the $n \times n$ identity and $L_\pm$ are $n \times n$ Hermitian matrices. Note that in this case J is automatically invertible. This form occurs naturally in many applications, but in fact the general JL spectral problem (1.1) can be rewritten in canonical form by an appropriate inflation of the dimension of the problem [37]. Under the assumption that $\mathrm{Ker}(L_+) \perp \mathrm{Ker}(L_-)$, a lower bound for the number of real points $\nu \in \sigma(JL)$ can be given in terms of the difference of the negative indices of operators $M_\pm$ that are suitable projections of the operators $L_\pm$.

THEOREM 5.6 (Lower bound for real points of $\sigma(JL)$). *Let L and J be $2n \times 2n$ matrices having canonical form* (5.14). *Suppose also that* $\mathrm{Ker}(L_+) \perp \mathrm{Ker}(L_-)$, *and let V denote the orthogonal complement of* $\mathrm{Ker}(L_+) \oplus \mathrm{Ker}(L_-)$ *in $Y = \mathbb{C}^n$ with corresponding orthogonal projection $P : Y \to V$. If $N_{\mathbb{R}}(JL) \leq 2n_{\mathrm{u}}$ denotes the number of nonzero real points in $\sigma(JL)$ counted with geometric multiplicity (necessarily an even number by the basic Hamiltonian symmetry $\sigma(JL) = -\overline{\sigma(JL)}$), then*

$$\frac{1}{2} N_{\mathbb{R}}(JL) \geq \left| N_-(M_+) - N_-(M_-) \right|, \qquad M_\pm := PL_\pm P. \tag{5.15}$$

Proof. First, we reduce the spectral problem $JLu = \nu u$ on $X = \mathbb{C}^{2n}$ to a spectral problem for a linear selfadjoint pencil on Y using the projection technique [24, 36, 37, 72]. Identifying $u \in X$ with the pair (u_1, u_2) with $u_j \in Y$ for $j = 1, 2$, the spectral problem $JLu = \nu u$ takes the form of a coupled system:

$$L_- u_2 = \nu u_1 \quad \text{and} \quad -L_+ u_1 = \nu u_2. \tag{5.16}$$

Now since $\mathrm{Ker}(L_+) \perp \mathrm{Ker}(L_-)$, the space Y can be decomposed into a direct sum of three pairwise orthogonal summands: $Y = \mathrm{Ker}(L_+) \oplus \mathrm{Ker}(L_-) \oplus V$. Letting $P_\pm : Y \to \mathrm{Ker}(L_\pm)$ be the orthogonal projections onto $\mathrm{Ker}(L_\pm)$, and using the facts that $L_\pm P_\pm = 0$ and (as $L_\pm$ are selfadjoint) $P_\pm L_\pm = 0$, we may apply the projections P_- and P to the equation $L_- u_2 = \nu u_1$ and hence obtain

$$\begin{aligned} 0 &= \nu P_- u_1 \,, \\ PL_- P_+ u_2 + PL_- P u_2 &= \nu P u_1 \,, \end{aligned} \tag{5.17}$$

and similarly applying the projections P_+ and P to the equation $-L_+ u_1 = \nu u_2$,

$$\begin{aligned} 0 &= \nu P_+ u_2 \,, \\ -PL_+ P_- u_1 - PL_+ P u_1 &= \nu P u_2 \,. \end{aligned} \tag{5.18}$$

If $\nu \neq 0$, then we obviously have $P_- u_1 = P_+ u_2 = 0$, and setting $v_j = P u_j \in V$ gives a coupled system of equations on the subspace V:

$$M_- v_2 = \nu v_1 \quad \text{and} \quad -M_+ v_1 = \nu v_2. \tag{5.19}$$

The advantage of this reduction is that $M_\pm$ are both invertible on the subspace V, and therefore by eliminating v_2 we can write the original spectral problem $JLu = \nu u$ for $\nu \neq 0$ as a spectral problem $\mathcal{L}(\lambda)v_1 = 0$ where $\mathcal{L}$ is the linear selfadjoint matrix pencil on V given by

$$\mathcal{L}(\lambda) := L_0 + \lambda L_1, \quad L_0 := M_+, \quad L_1 := M_-^{-1}, \quad \lambda := \nu^2. \tag{5.20}$$

Now we invoke Theorem 5.2, specifically the inequality (5.2) for the number $N_+(\mathcal{L})$ of positive real characteristic values of $\mathcal{L}$ counted with geometric multiplicity. Since L_0 is invertible as well as the leading coefficient L_1, and since the total number of curves $\mu = \mu(\lambda)$ associated with the pencil $\mathcal{L}$ passing through the origin $(\lambda, \mu) = (0, 0)$ is equal to $\dim(\mathrm{Ker}(L_0))$, there are no such curves at all in this application, thus $Z_\pm^\downarrow(\mathcal{L}) = 0$. Therefore (5.2) takes the simple form

$$N_+(\mathcal{L}) \geq |N_-(L_0) - N_-(L_1)| = \left|N_-(M_+) - N_-(M_-^{-1})\right| = |N_-(M_+) - N_-(M_-)|\,. \tag{5.21}$$

The inequality (5.15) follows from the fact that $N_+(\mathcal{L})$ counts the number of positive points in $\sigma(\mathcal{L})$, and hence by $\lambda = \nu^2$ it is exactly half the number of real nonzero points in $\sigma(JL)$, counted with geometric multiplicity. □

5.3. Historical remarks. Corollary 5.5 is a generalization of an index theorem proved independently by Kapitula *et al.* [37] and Pelinovsky [65] in the special case that each maximal chain of root vectors corresponding to the characteristic value $\lambda = 0$ has length 2. Such a situation corresponds to many typical Hamiltonian systems with symmetries [26, 37]. In this case ζ can be written in the form $N_-(D)$ where D is a Gram-type Hermitian matrix with elements $D_{jk} := (Lu^{[j]}, u^{[k]})$ and the vectors $u^{[j]}$ span $\mathrm{gKer}(JL) \ominus \mathrm{Ker}(L)$.

The relation between the dimension $2n_\mathrm{u}$ of the unstable invariant subspace of JL and the number of negative eigenvalues of L was first studied for $N_-(L) = 1$, in which case under the assumption that $\dim(\mathrm{gKer}(JL)) = 2\dim(\mathrm{Ker}(L))$, the count (5.12) indicates that JL has at most one pair of non-imaginary (unstable) points of spectrum, and by the full Hamiltonian symmetry these points are necessarily real. In this case the famous *Vakhitov-Kolkolov criterion* [70] applies in a large class of problems in the stability theory of waves and identifies the scalar quantity $\zeta = Z^+(\mathcal{L}) = N_-(D)$ that determines whether $n_\mathrm{u} = 0$ or $n_\mathrm{u} = 1$, i.e., whether a given wave is spectrally stable. In the Vakhitov-Kolokolov theory, the quantity $N_-(D)$ turns out to be equal to the derivative of the momentum (impulse functional) of the wave profile with respect to the wave speed. Pego and Weinstein [64] proved that the same quantity is related to the second derivative $D''(0)$ of the (usual) Evans function $D(\lambda)$, which by symmetries of the system satisfies $D(0) = D'(0) = 0$. In celebrated papers, Grillakis *et al.* [26, 27] extended the spectral stability analysis associated to the Vakhitov-Kolokolov instability criterion to establish fully nonlinear stability properties of waves. See [5] for the analysis preceding these general results and [38] for a historical discussion. Simultaneously, Maddocks [56] developed the theory of inertial laws for constrained Hamiltonians and related it to results of MacKay [55] giving rise to an algebraic method for proving index theorems.

Kapitula *et al.* [37] and Pelinovsky [65] proved a generalization of the Vakhitov-Kolokolov criterion to the case of $N_-(L) > 1$. They derived the formula mentioned above for $\zeta = Z^+(\mathcal{L})$ in terms of the Gram matrix D, and they also interpreted the law (5.12) as a parity index by writing it in the form $n_\mathrm{u} \equiv N_-(L) - N_-(D) \pmod 2$. This identity plays an important role in the presence of full Hamiltonian symmetry

because oddness of n_{u} indicates the presence of purely real points of the unstable spectrum. Under some technical assumptions, Kapitula [36] reproved these results using a similar technique involving continuity of eigenvalue curves in conjunction with the so-called *Krein matrix* method. The latter effectively projects the problem to a finite-dimensional negative subspace of L, but in the process introduces unnecessary poles in the eigenvalue branches $\mu(\lambda)$, and these poles obstruct the simple graphical visualization of the Krein signatures and indices arising in the analysis. Kapitula and Haragus [29] also proved the same count in the setting of differential operators with periodic coefficients using the Bloch wave decomposition (Floquet theory) to reduce the problem with bands of continuous spectrum to a collection of problems each of which has purely discrete spectrum. See also [8] for an alternative proof of (5.12) using aspects of integrable systems theory. An analogous index theorem for quadratic Hermitian matrix pencils was proved by Pelinovsky and Chugunova [10] using the theory of indefinite quadratic forms (Pontryagin spaces) and was later reproved by a homotopy technique [43] resembling the graphical method described in this paper. Another related count for quadratic operator pencils was recently proved by Bronski *et al.* [9]. Theorem 5.2 is an example of the kind of results that can be obtained beyond the simple context of problems reducible to linear pencils.

When the reality condition guaranteeing full Hamiltonian symmetry is dropped and one has instead of (5.12) the more general statement (5.6), it is no longer possible to interpret the inertia law as a parity index. Results analogous to those recorded in Theorem 5.3 were previously proven in [11, Theorem 6] and in [29, Theorem 2.13] under the assumption that L is an invertible operator with a compact inverse. For the reader familiar with [11, 29], the apparent difference between (5.6) and the formulation of the index theorems found therein is caused by a different choice of the definition of the Krein signature; see the footnote referenced between equations (1.7) and (1.8).

Theorem 5.6, which concerns problems having the canonical symplectic structure (5.14), was proved independently and virtually simultaneously by Jones [34] and Grillakis [24]. Since the inequality (5.15) only provides a lower bound on the number of purely real points of $\sigma(JL)$ and not its exact count, the straightforward generalization from the finite-dimensional setting to operator theory does not require completeness of the root vectors. Jones' proof [34] of Theorem 5.6 is of a very different nature from the graphical one we have presented, but that of Grillakis [24] is quite similar, with spectral projections playing the role of the eigenvalue curves $\mu = \mu(\lambda)$. We think that our approach, embodied in the proof of Theorem 5.2, gives a very simple way to visualize the count. This problem was also studied by Chugunova and Pelinovsky [11], who proved a number of related results using the Pontryagin Invariant Subspace Theorem applied to the linear pencil. More recently, a similar approach combined with the use of the Krein matrix, an analytical interpretation of the Krein signature, and the Keldysh Theorem [40, 57], was employed by Kapitula and Promislow [38], who reproved Theorem 5.6 (see the paper for further historical remarks). The connection to the linear pencil was also pointed out in a similar setting in [72].

6. Discussion and Open Problems. We hope that our paper has demonstrated how the analytic interpretation of Krein signatures and indices in terms of real curves and their order of vanishing at real characteristic values helps to easily visualize, simplify, and organize numerous results found in the literature on stability of nonlinear waves.

The analytical or graphical interpretation of the Krein signature put forth in Definition 3.3 and Theorem 3.4 is apparently limited to real characteristic values

of selfadjoint pencils. In particular, such a formulation applied to the linear pencil (1.4) related to the Hamiltonian spectral problem (1.1) does not provide any direct information about $\sigma(JL)$ off the imaginary axis. But such a characterization seems unnecessary, as the (traditional) Krein signature of non-imaginary points of $\sigma(JL)$ is easy to calculate and is equal to zero. Nevertheless it would be interesting to determine whether an approach similar to the one presented here can be applied to study non-selfadjoint pencils and/or to detect further information about non-real characteristic values of selfadjoint pencils.

Our main new result, the generalization of the notion of an Evans function to that of an Evans-Krein function, simplifies numerical calculation of Krein signatures from an Evans function and is a method easy to incorporate into existing codes. From the theoretical perspective an interesting problem would be to find an intrinsically geometrical interpretation of Evans-Krein functions similar to the characterization of Evans functions given by Alexander *et al.* [1].

The graphical nature of the signature should also allow us to generalize the index theorems presented in §5 to handle operators on infinite-dimensional Hilbert spaces with general kernels, sidestepping certain unnecessary technical difficulties. One also expects to be able to identify optimal (from the graphical point of view) assumptions for the validity of these theorems. We advocate that, in general, the concept of the graphical signature is often more suitable for analysis (either rigorous or numerical) than the traditional one as it does not rely on any particular algebraic structure of the operator pencil. This may possibly allow further applications of signature arguments for non-polynomial pencils where the traditional definition based on indefinite quadratic forms falls flat.

Furthermore, the graphical approach to Krein signature theory may be preferable for understanding various mechanisms for avoiding Hamiltonian-Hopf bifurcations [55, 58] in Hamiltonian systems with a tuning parameter $t \in \mathbb{R}$. Of particular interest here are "non-generic" collisions (in the sense of Arnold [2]) involving two real characteristic values of a selfadjoint pencil having opposite Krein signatures but that nonetheless do not bifurcate from the real axis. The question to be addressed is the fate of the various eigenvalue curves $\mu = \mu(\lambda)$ of (1.5) when the spectral problem is perturbed in an admissible fashion. Here, one expects that certain conditions may ensure that not only do real characteristic values of opposite signatures survive collisions, but also the transversal intersection of eigenvalue branches $\mu = \mu(\lambda)$ is preserved as well. Such a mechanism for avoiding Hamiltonian-Hopf bifurcations is significantly different from the one described by Krein and Ljubarskii [49] in the context of differential equations with periodic coefficients. Here we only present an illustrative example of what we have in mind; we expect to publish elsewhere new results given necessary and sufficient conditions for the preservation of branch crossings.

$\lhd$ *Example 14. Nongeneric perturbations can avoid Hamiltonian-Hopf bifurcations.* Consider the pencil $\mathcal{L}(\lambda) := L - \lambda K$, $K = (iJ)^{-1}$, where L depends linearly on a control parameter $t \in \mathbb{R}$ as $L = A + tB$ where $A := \operatorname{diag}(5, 1, 3, 2)$, and where

$$B := \begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \quad \text{and} \quad J := -\frac{1}{2} \begin{pmatrix} 0 & 2i & 0 & 0 \\ 2i & 0 & 0 & 0 \\ 0 & 0 & 0 & i \\ 0 & 0 & i & 0 \end{pmatrix}. \tag{6.1}$$

Fig. 6.1 illustrates how transversal intersections between the eigenvalue branches $\mu = \mu(\lambda)$ present for $t = 0$ can persist for $t \neq 0$. For general perturbations of

$L = A$ one would expect all four transversal intersections to simultaneously "break" for $t \neq 0$, resulting in four uniformly ordered analytic eigenvalue branches $\mu_1(\lambda) < \mu_2(\lambda) < \mu_3(\lambda) < \mu_4(\lambda)$, and if these branches have critical points, Hamiltonian-Hopf bifurcations resulting in the loss of one or more pairs of real characteristic values becomes possible as the branches shift vertically. Therefore the perturbation B has to be quite special for the intersections to persist, a situation that prevents all Hamiltonian-Hopf bifurcations from occurring. ▷

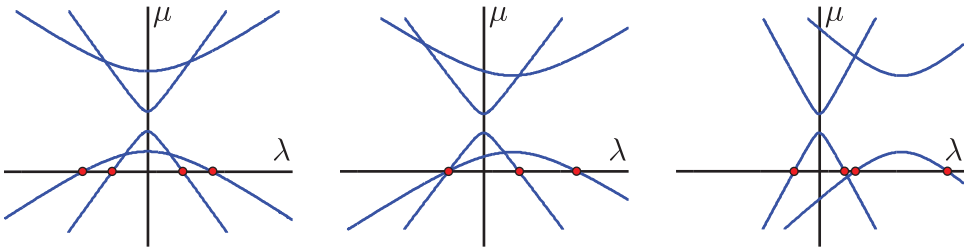

FIG. 6.1. *Spectrum of $L(t) - \lambda K$ plotted against λ for the matrices in Example 14. Left panel: $t = 0$. Center panel: $t = 1$ (exhibiting a necessarily harmless collision between characteristic values of the same Krein signature). Right panel: $t = 4$. Between $t = 1$ and $t = 4$ two real characteristic values of opposite Krein signature pass through each other unscathed due to the preservation of transversal crossings of eigenvalue branches $\mu = \mu(\lambda)$.*

Finally, we would like to stress that mostly for the sake of keeping our presentation as simple and accessible as possible, we have chosen to work in finite-dimensional spaces (except in Sections 2 and 4). Extending the theory to operators on infinite-dimensional spaces and general self-adjoint operator pencils is a task that has both straightforward aspects and also technically difficult aspects. One of the central open questions in the field of stability of nonlinear waves is whether it is possible to prove an analogue of Corollary 5.5 in the case of an unbounded operator J, as in the important case $J = \partial_x$ relevant to the stability of solutions to the Korteweg-de Vries equation [64]. The analysis presented here suggests that once the key assumption of continuity of eigenvalue branches of the pencil $\mathcal{L}(\lambda)$ is met, the theory developed for bounded operators can be generalized to unbounded operators. Our results in this direction will be published elsewhere.

REFERENCES

[1] J. ALEXANDER, R. GARDNER, C. K. R. T. JONES, *A topological invariant arising in the stability analysis of traveling waves*, J. Reine Angew. Math. 410 (1990), pp. 167–212.
[2] V. I. ARNOLD, *Mathematical methods of classical mechanics*, Springer-Verlag, Heidelberg-Berlin, 1978.
[3] V. I. ARNOLD, A. AVEZ, *Ergodic Problems of Classical Mechanics*, Benjamin, New York, 1968.
[4] T. BETCKE, N. J. HIGHAM, V. MEHRMANN, C. SCHRÖDER, AND F. TISSEUR, *NLEVP: A Collection of Nonlinear Eigenvalue Problems*, MIMS EPrint 2011.116, December 2011.
[5] J. L. BONA, P. E. SOUGANIDIS, W. A. STRAUSS, *Stability and instability of solitary waves of KdV type*, Proc. R. Soc. Lond. A 411 (1987), pp. 395–412.
[6] T. J. BRIDGES, G. DERKS, *The Symplectic Evans matrix, and the instability of solitary waves and fronts*, Arch. Rat. Mech. Anal. 156 (2001), pp. 1–87.
[7] J. C. BRONSKI, M. A. JOHNSON, Krein Signatures for the Faddeev-Takhtajan Eigenvalue Problem, Comm. Math. Phys. 288 (2009), no. 3, pp. 821–846.

[8] J. C. Bronski, M. A. Johnson, and T. Kapitula, *An index theorem for the stability of periodic traveling waves of KdV type*, Proc. Roy. Soc. Edinburgh: Section A 141 (2011), pp. 1141–1173.
[9] J. C. Bronski, M. A. Johnson, and T. Kapitula, *An instability index theory for quadratic pencils and applications*, preprint, 2012.
[10] M. Chugunova, D. Pelinovsky, *On quadratic eigenvalue problems arising in stability of discrete vortices*, Linear Algebra and its Applications, 431 (2009), pp. 962–973.
[11] M. Chugunova, D. Pelinovsky, *Count of eigenvalues in the generalized eigenvalue problem*, J. Math. Phys. 51 (2010), pp. 052901.
[12] J. C. DiFranco and P. D. Miller, *The semiclassical modified nonlinear Schrd̈ingier equation I: modulation theory and spectral analysis*, Physica D, 237 (2008), pp. 947–997.
[13] P. G. Drazin, W. H. Reid, *Hydrodynamic Stability*, Cambridge University Press, 2004.
[14] C. Effenberger, D. Kressner, and C. Engström, *Linearization techniques for band structure calculations in absorbing photonic crystals*, Int. J. Numer. Meth. Eng. 89 (2012), pp. 180–191.
[15] J. W. Evans, *Nerve axon equations. I. Linear approximations*, Indiana Univ. Math. J. 21 (1972), pp. 875–955.
[16] J. W. Evans, *Nerve axon equations. II. Stability at rest*, Indiana Univ. Math. J. 22 (1972), pp. 75–90.
[17] J. W. Evans, *Nerve axon equations. III. Stability of the nerve impulse*, Indiana Univ. Math. J. 22 (1972), pp. 577–594.
[18] J. W. Evans, *Nerve axon equations. IV. The stable and unstable impulse*, Indiana Univ. Math. J. 24 (1975), pp. 1169–1190.
[19] F. Gesztesy, Y. Latushkin, K. A. Makarov, *Evans functions, Jost functions, and Fredholm determinants*, Arch. Rat. Mech. Anal. 186 (2007), pp. 361–421.
[20] I. C. Gohberg, M. G. Krein, *Introduction to the theory of linear nonselfadjoint operators*. Transactions Mathematical Monographs, no. 18. Providence, RI: American Mathematical Society, 1969.
[21] I. Gohberg, P. Lancaster, L. Rodman, *Matrix polynomials*. Academic Press New York, 1982. Reprinted by SIAM Philadelphia, 2009.
[22] I. Gohberg, P. Lancaster, L. Rodman, *Invariant Subspaces of Matrices with Applications*. John Wiley & Sons, Inc., New York, 1986. Reprinted by SIAM, Philadelphia, 2006.
[23] I. Gohberg, P. Lancaster, L. Rodman, *Indefinite Linear Algebra and Applications*, Basel: Birkhäuser Verlag, 2005.
[24] M. Grillakis, *Linearized instability for nonlinear Schrödinger and Klein-Gordon equations*, Comm. Pure Appl. Math. 46 (1988), pp. 747–774.
[25] M. Grillakis, *Analysis of the linearization around a critical point of an infinite dimensional Hamiltonian system*, Communs Pure Appl. Math 43 (1990), pp. 299–333.
[26] M. Grillakis, J. Shatah, W. A. Strauss, *Stability theory of solitary waves in the presence of symmetry I*, J. Funct. Analysis 74 (1987), pp. 160–197.
[27] M. Grillakis, J. Shatah, W. A. Strauss, *Stability theory of solitary waves in the presence of symmetry II*, J. Funct. Analysis 94 (1987), pp. 308–348.
[28] K. F. Gurski, R. Kollár, R. L. Pego, *Slow damping of internal waves in a stably stratified fluid*, Proc. R. Soc. Lond. A 460 (2004), pp. 977–994.
[29] M. Haragus and T. Kapitula, *On the spectra of periodic waves for infinite-dimensional Hamiltonian systems*, Physica D 237 (2008), pp. 2649–2671.
[30] R. O. Hryniv, P. Lancaster, A. A. Renshaw, *A stability criterion for parameter-dependent gyroscopic systems*, Trans. ASME J. Appl. Mech. 66 (1999), pp. 660–664.
[31] J. Humpherys, B. Sandstede and K. Zumbrun, *Efficient computation of analytic bases in Evans function analysis of large systems*, Numerische Mathematik 103 (2006), pp. 631–642.
[32] I. S. Iohvidov, M. G. Krein, *Spectral theory of operators in spaces with indefinite metric I*, Tr. Mosk. Mat. Obsc 5 (1956), pp. 367–432 (Transl. Amer. Math. Soc. Transl. 2 (1960), pp. 105–175).
[33] E. Jarlebring, K. Meerbergen, W. Michiels, *A Krylov method for the delay eigenvalue problem*, SIAM J. Sci. Comput. 32 (2010), pp. 3278–3300.
[34] C. K. R. T. Jones, *Instability of standing waves for non-linear Schrödinger-type equations*, Ergod. Th. & Dynam. Sys. 8 (1988), pp. 119–138.
[35] C. K. R. T. Jones, R. Marangell, P. D. Miller, and R. G. Plaza, *The spectrum of periodic wave trains for the nonlinear Klein-Gordon equation*, in preparation. (2012).
[36] T. Kapitula, *The Krein signature, Krein eigenvalues, and the Krein Oscillation Theorem*, Indiana U. Math. J. 59 (2010), pp. 1245–1276.
[37] T. Kapitula, P. Kevrekidis, and B. Sandstede, *Counting eigenvalues via the Krein sig-*

nature in infinite-dimensional Hamiltonian systems, Physica D, 195 (2004), pp. 263–282; Addendum: Physica D, 201 (2005), pp. 199–201.
[38] T. Kapitula, K. Promislow, *Stability indices for constrained self-adjoint operators*, Proc. Amer. Math. Soc. 140 (2012), pp. 865–880.
[39] T. Kato, *Perturbation Theory for Linear Operators*, Springer Verlag, 1976.
[40] M. V. Keldyš, *On the characteristic values and characteristic functions of certain classes of non-self-adjoint equations*, Dokl. Akad. Nauk SSSR 77 (1951), pp. 11–14.
[41] O. N. Kirillov, U. Günther, F. Stefani, *Determining role of Krein signature for three-dimensional Arnold tongues of oscillatory dynamos*, Phys. Rev. E 79 (2009), 016205.
[42] M. Klaus, J. K. Shaw, *On the eigenvalues of Zakharov-Shabat systems*, SIAM J. Math. Anal. 34 (2003), No. 4, pp. 759–773.
[43] R. Kollár, *Homotopy method for nonlinear eigenvalue pencils with applications*, SIAM J. Math. Anal. 43 (2011), pp. 612–633.
[44] R. Kollár, R. L. Pego, *Spectral stability of vortices in two-dimensional Bose-Einstein condensates via the Evans function and Krein signature*, Appl. Math. Res. Express 2012 (2012), pp. 1–46.
[45] M. G. Krein, *A generalization of some investigations of A. M. Lyapunov on linear differential equations with periodic coefficients*, Doklady Akademii Nauk SSSR 73 (1950), pp. 445–448.
[46] M. G. Krein, *On the application of an algebraic proposition in the theory of matrices of monodromy*, Uspekhi Mat. Nauk 6:1(41) (1951), pp. 171–177.
[47] M. G. Krein, H. Langer, *On some mathematical principles in the linear theory of damped oscillations of continua I*, Integral Equations and Operator Theory 1 (1978), pp. 364–399.
[48] M. G. Krein, H. Langer, *On some mathematical principles in the linear theory of damped oscillations of continua II*. Integral Equations and Operator Theory 1 (1978), pp. 539–566.
[49] M. G. Krein, G. J. Ljubarskii, *Analytic properties of multipliers of periodic canonical differential systems of positive type*, Amer. Math. Soc. Transl. 89 (1970), pp. 1–28 (Translation of Izv. Akad. Nauk SSSR Ser. Mat. 26 (1962), pp. 549–572).
[50] P. Lancaster, *Lambda-matrices and Vibrating Systems*, Oxford: Pergamon Press, 1966.
[51] P. Lancaster, A. S. Markus, F. Zhou, *Perturbation theory for analytic matrix functions: The semisimple case*, SIAM J. Matrix Anal. Appl. 25 (2003), No. 3, pp. 606–626.
[52] P. Lancaster, M. Tismenetsky, *The Theory of Matrices*, 2nd ed., London: Academic Press, 1985.
[53] P. Lancaster, P. Zizler, *On the stability of gyroscopic systems*, J Appl. Mech. 65 (1998), pp. 519–522.
[54] Y. Li and K. Promislow, *Structural stability of non-ground state traveling waves of coupled nonlinear Schrödinger equations*, Physica D 124 (1998), no. 1–3, pp. 137–165.
[55] R. S. MacKay, *Stability of equilibria of Hamiltonian systems.* In: Hamiltonian dynamical systems (ed. R. S. MacKay & J. D. Meiss), pp. 137–153. Bristol: Hilger.
[56] J. H. Maddocks, *Restricted quadratic forms and their application to bifurcation and stability in constrained variational principles*, SIAM J. Math. Anal. 16 (1988), pp. 47–68; Errata 19 (1988), pp. 1256–1257.
[57] A. S. Markus, *Introduction to the Spectral Theory of Polynomial Operator Pencils*, Translations of Mathematical Monographs, no. 71. Providence, RI: American Mathematical Society, 1988.
[58] J. D. Meiss, *Differential dynamical systems*, SIAM Philadelphia 2007.
[59] M. Sh. Misrikhanov, V. N. Ryabchenko, *The quadratic eigenvalue problem in electric power systems*, Automation and Rem. Control 67 (2006), no. 5, pp. 698–720.
[60] V. Mehrmann, H Voss, *Nonlinear eigenvalue problems: A challenge for modern eigenvalue methods*, GAMM-Mitteilungen (GAMM-Reports) 27 (2004), pp. 121–152.
[61] T. S. Motzkin, O. Taussky, *Pairs of matrices with property L*, Trans. Am. Math. Soc. 73 (1952), pp. 108–114.
[62] T. S. Motzkin, O. Taussky, *Pairs of matrices with property L. II.* Trans. Am. Math. Soc. 80 (1955), pp. 387–401.
[63] R. L. Pego, H. Warchall, *Spectrally stable encapsulated vortices for nonlinear Schrödinger equations*, J. Nonlinear Sci. 12 (2002), pp. 347–394.
[64] R. L. Pego, M. I. Weinstein, *Eigenvalues, and instabilities of solitary waves*, Philos. Trans. Roy. Soc. London Ser. A 340 (1992), pp. 47–94.
[65] D. E. Pelinovsky, *Inertia law for spectral stability of solitary waves in coupled nonlinear Schrödinger equations*, Proc. Roy. Soc. Lond. A 461 (2005), pp. 783–812.
[66] L. S. Pontryagin, *Hermitian operators in spaces with indefinite metric*, Izv. Akad. Nauk SSSR Ser. Mat. 8 (1944), pp. 243–280.
[67] F. Rellich, *Störungstheorie der Spektralzerlegung. V*, Math. Ann. 118 (1942), pp. 462–484.

[68] F. Rellich, *Perturbation theory of eigenvalue problems*, New York University, 1953.
[69] F. Tisseur, K. Meerbergen, *The quadratic eigenvalue problem*, SIAM Review 43 (2001), pp. 235–286.
[70] N. G. Vakhitov, A. A. Kolokolov, *Stationary solutions of the wave equation in a medium with nonlinearity saturation*, Izv. Vyssh. Uchebn. Zaved. Radiofiz. 16 (1973), pp. 1020–1028 (Radiophys. Quantum Electron. 16 (1973), pp. 783–789).
[71] J.-C. van der Meer, *The Hamiltonian Hopf Bifurcation*. Springer-Verlag, Berlin, 1985.
[72] V. Vougalter, D. Pelinovsky, *Eigenvalues of zero energy in the linearized NLS problem*, J Math Phys 47 (2006), 062701.
[73] M. Wadati, K. Konno, and Y. Ichikawa, *Generalization of inverse scattering method*, J. Phys. Soc. Japan 46 (1979), pp. 1965–1966.
[74] V. A. Yakubovitch, V. M. Starzhinskii, *Linear Differential Equations with Periodic Coefficients*. John Wiley and Sons, New York, 1975.
[75] E. Yanagida, *Stability of fast traveling pulse solutions of the FitzHugh-Nagumo equations*, J. math. Biol. 22 (1985), pp. 81–104.
[76] V. E. Zakharov and A. B. Shabat, *Exact theory of two-dimensional self-focusing and one-dimensional self-modulation of waves in nonlinear media*, Sov. Phys. JETP 34 (1972), pp. 62–69.